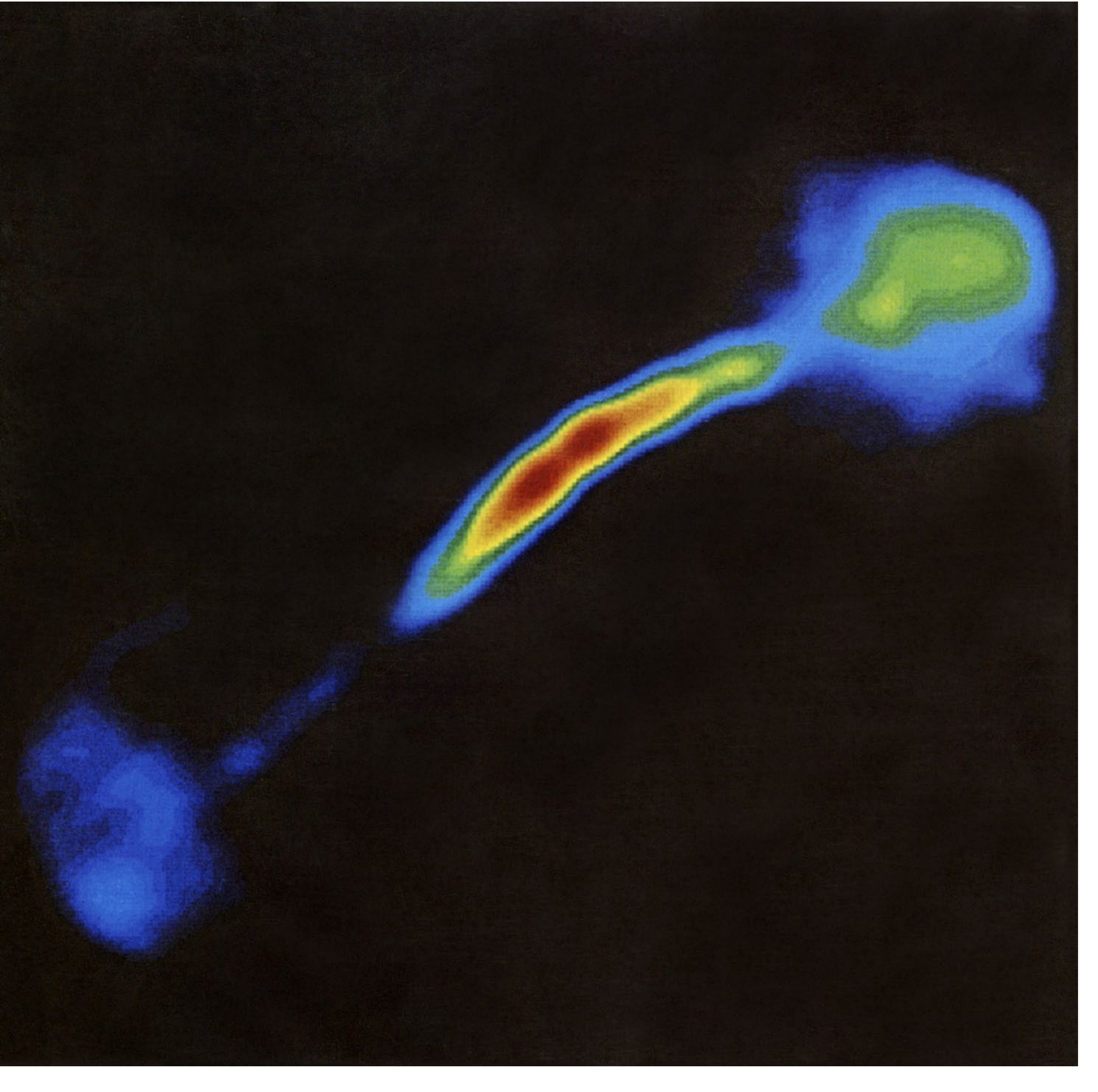
INVESTIGACION Y CIENCIA

Edición en español de **SCIENTIFIC AMERICAN**



CHORROS COSMICOS

Julio 1982
300 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 70
Julio 1982

INVESTIGACION Y
CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

EDITA Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

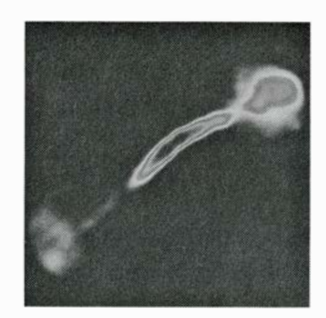
Alberto Broggi y Víctor Obiols: *Esperanza de vida y crecimiento demográfico en el Tercer Mundo*; Antonio Méndez Vilaseca: *Quarkonio*; Emilio Gelpí: *Imágenes por RMN en medicina*; Santiago Torres Martínez: *Bases genéticas de la diversidad de anticuerpos*; J. M. Alonso-Allende Yohn: *La asimetría de las platijas*; Manuel Puigcerver: *Chorros cósmicos*; Julio Samsó Moya: *El calendario gregoriano*; Luis Bou: *Temas metamágicos*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Ramón Margalef, Ramón Canal, Jordi Isern y Javier Labay

Libros:

Antonio Escarré, Oriol de Bolós, Mariano Artigas y Claudi Esteva



LA PORTADA

La ilustración de la portada muestra la configuración de la radioemisión procedente de la galaxia elíptica gigante IC 4296, cartografiada con ayuda de la Gran Distribución (VLA), conjunto de 27 telescopios interconectados en un enclave al Oeste de Socorro, Nuevo México. La galaxia, probablemente miembro del cúmulo de galaxias del Centauro, se encuentra a una distancia de 120 millones de años-luz. A pesar de lo cual se trata de la radiofuente "clásica" más cercana. Esas fuentes constan de dos lóbulos simétricos de radioemisión. Los primitivos radiomapas de IC 4296 mostraban una mancha de radioemisión entre los lóbulos. Tres años más tarde, el trabajo con el VLA reveló que la mancha estaba constituida por dos chorros de gas que salían en direcciones opuestas del centro de la galaxia. El mapa.de la portada se obtuvo de observaciones hechas en noviembre. Los chorros aparecen en amarillo y rojo, los lóbulos en verde y azul. Entre lóbulos media una distancia de casi un millón de años-luz.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:
España:
Un año (12 números): 3.300 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
340 pesetas
Ejemplar atrasado extraordinario:
440 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.°, Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

La portada es cortesía del
Observatorio Radioastronómico
Nacional de los Estados Unidos

Página	Fuente
9-16	Albert E. Miller
19-30	Gabor Kiss
33	Technicare Corporation
34-40	Walken Graphics
41	Technicare Corporation (*arriba y en medio*). *Journal of Computerized Axial Tomography* (*abajo*)
42	James M. S. Hutchison, Francis Smith y John R. Mallard (*izquierda y centro*), Ilil Arbel (*derecha*)
43	Walken Graphics
51	Richard J. Feldman, National Institutes of Health (*arriba*); Bunji Tagawa (*abajo*)
52-55	Bunji Tagawa
56	Ilil Arbel (*arriba y derecha*); Philip Leder, Harvard Medical School (*izquierda*)
57-61	Bunji Tagawa
67	Albert M. Plaza
68-72	Tom Prentiss

Página	Fuente
75-78	National Radio Astronomy Observatory
79-84	Ian Worpole
87	Archivos de la Ciudad de Siena
88	Gordon Moyer
89	Biblioteca del Congreso
90	Edmondo Lamalle, S. J. (*izquierda*), Biblioteca de la Universidad de Cornell (*derecha*)
91	Biblioteca Nacional de Nápoles, fotografía de Giovanna De Vita (*izquierda*); Biblioteca Nacional Central de Florencia, fotografía de Guido Sansoni (*derecha*)
92-93	Ilil Arbel
94	Fernando de Castro (*arriba*), G. Muratori (*abajo*)
96	Carlos Eyzaguirre y Patricio Zapata
97-108	Carlos Eyzaguirre, Patricio Zapata y Miguel Alonso
111	Ray Atkeson
116-120	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

DAVIDSON R. GWATKIN y SARAH K. BRANDEL ("Esperanza de vida y crecimiento demográfico en el Tercer Mundo") se hallan adscritos al Consejo del Desarrollo de Ultramar de Washington. Colaboran en un estudio del índice de mortalidad de los países en vías de desarrollo y del efecto de la ayuda internacional sobre el crecimiento demográfico. Gwatkin se licenció por el Haverford College en 1962. Cursó el segundo ciclo universitario en Princeton. Desde 1964 hasta 1977 trabajó para la Fundación Ford. Brandel se licenció por la Universidad de Denison en 1963. Se doctoró en estudios sudasiáticos por la Universidad de California en Berkeley (1965) y en demografía y sociología por la de Georgetown (1977). En 1978 trabajó en el servicio de investigación del Banco Mundial; ese mismo año entró en el Consejo del Desarrollo de Ultramar. Gwatkin y Brandel agradecen la ayuda prestada por la Oficina Norteamericana para el Desarrollo Internacional y la del Fondo de las Naciones Unidas para Actividades de la Población a la hora de recabar los datos en que fundan su artículo.

ELLIOTT D. BLOOM y GARY J. FELDMAN ("Quarkonio") enseñan física en la Universidad de Stanford. Ambos han venido trabajando sobre los estados de materia creados en aniquilaciones electrón-positrón. Bloom se licenció en 1962 por el Pomona College y se doctoró, en 1967, por el Instituto de Tecnología de California. En 1967 se trasladó a Stanford como investigador asociado, subiendo desde entonces todos los peldaños del escalafón docente. Ha colaborado con un vasto grupo de físicos en Europa y los Estados Unidos en experimentos de interacciones electrón-positrón. Feldman cursó el primer ciclo en la Universidad de Chicago; se licenció (1965) y doctoró (1971) por la Universidad de Harvard. Ingresó en Stanford en 1971 como investigador adjunto; en 1979 se le nombró profesor adjunto.

IAN L. PYKETT ("Imágenes por RMN en medicina") disfruta de una beca de investigación en física en el Hospital General de Massachusetts. Licenciado en 1974 por el Politécnico de Londres, se doctoró en física, en 1978, por la Universidad de Nottingham. Desde 1978 hasta 1980 ocupó una plaza de investigador en Nottingham, trabajando en técnicas sobre formación de imágenes. En 1980 se trasladó a los Estados Unidos para incorporarse a la Technicare Corporation.

PHILIP LEDER ("Bases genéticas de la diversidad de anticuerpos") es titular de la cátedra de genética John Emory Andrus y director del departamento de genética de la Facultad de Medicina de Harvard. Se licenció por el Harvard College, en 1956, doctorándose en medicina en 1960. En 1962 ingresó como investigador adjunto en el Instituto Nacional del Corazón. Tras sendos períodos de investigador médico en el Instituto Nacional del Cáncer, jefe de la sección de genética molecular del Instituto Nacional de la Salud Infantil y el Desarrollo Humano (NICHD) y jefe del laboratorio de genética molecular del NICHD, se trasladó, en 1980, a Harvard.

DAVID POLICANSKY ("La asimetría de las platijas") trabaja en el Gray Herbarium de la Universidad de Harvard. Nacido en Sudáfrica, cursó sus estudios de liceo en su país natal y en Inglaterra, y los universitarios en los Estados Unidos, en Stanford. En 1972 se doctoró por la Universidad de Oregón. Antes de ingresar en Harvard enseñó en la Universidad de Massachusetts en Boston. Sus investigaciones se centran en la evolución de la reproducción sexual y en el control temporal de distintos sucesos del ciclo vital, como la reproducción y la metamorfosis. Quiere dejar constancia del auxilio prestado en la obtención de platijas adultas y en el cultivo de larvas de platija por el personal del New England Acuarium de Boston, especialmente Louis Garibaldi, Michael Kelleher, Bruce Hecker y Paul Sieswerda; el personal del Acuario Steinhart de San Francisco, en particular John McCosker; Kunio Amaoka, de la Universidad de Hokkaido, Stanley Fullerton, de Santa Cruz, y Jeffery Meyer, de San Francisco.

ROGER D. BLANDFORD, MITCHELL C. BEGELMAN y MARTIN J. REES ("Chorros cósmicos") son astrofísicos con un interés común por las fuentes de radioondas situadas fuera de nuestra galaxia. Blandford es inglés de nacimiento, formado en la Universidad de Cambridge. Estuvo en el Magdalene College de Cambridge, pasando luego al St. John's College, también en Cambridge. Durante el bienio 1974-75 perteneció al Instituto de Estudios Superiores de la Universidad de Princeton. En 1976, marchó al Instituto de Tecnología de California, donde enseña física teórica. Begelman realiza investigación postdoctoral en la Universidad de California en Berkeley. Comenzó su carrera en Harvard, si bien su doctorado, concedido en 1978, es de la Universidad de Cambridge. Se trasladó a Berkeley en 1979. Al igual que Blandford, Rees es también inglés. Detenta la cátedra Plume de astronomía en Cambridge, amén de dirigir su Instituto de Astronomía.

GORDON MOYER ("El calendario gregoriano") es un escritor independiente dedicado a temas relacionados con la historia de la ciencia. Tiene especial interés en los principios sobre los que se fundamentan los calendarios. Se licenció en filología inglesa por la Universidad de Maryland. Hasta 1980 fue asesor del Planetario Davis de la Academia de Ciencias de Maryland en temas relacionados con el calendario y con la cronología de acontecimientos astronómicos. En el año antes citado se trasladó a Frankfurt am Main, Alemania Federal, donde dio conferencias sobre el calendario gregoriano en el Institut für Geschichte der Naturwissenschaften de la Universidad Goethe, por invitación del difunto profesor Willy Hartner.

CARLOS EYZAGUIRRE y PATRICIO ZAPATA ("El cuerpo carotídeo") son fisiólogos, cuya investigación se ha centrado durante los últimos años en los problemas del funcionamiento de quimiorreceptores arteriales. Ambos estudiaron en la Facultad de Medicina de la Universidad Católica de Chile, licenciándose por la Universidad de Chile, en 1947 y 1963, respectivamente. Eyzaguirre es profesor y jefe del departamento de fisiología de la Universidad de Utah. Discípulo de V. Luco (U. Católica de Chile) y S. W. Kuffler (U. Johns Hopkins), ha sido consultor del Instituto Nacional de la Salud estadounidense y es Doctor Honoris Causa por las Universidades Católica de Chile y Complutense de Madrid. Zapata se formó como fisiólogo junto a Eyzaguirre, en Utah, donde realizó tareas investigadoras (1966-1968) y docentes (1973-1974). Actualmente es profesor y jefe del laboratorio de neurobiología de la Universidad Católica de Chile.

Hace...

José M.ª López Piñero

... doscientos cincuenta años

Andrés Piquer Arrufat estudiaba medicina en la Universidad de Valencia. Había nacido en 1711 en la localidad turolense de Fórnoles, donde también había cursado las primeras letras.

Los profesores que más influyeron en la formación científica de Piquer fueron el catedrático de anatomía Juan Bautista Longás y, sobre todo, Antonio García Cervera, titular de la cátedra médica primaria y antes de la de práctica. Piquer lo llamó "García el grande" y lo consideró cabeza destacada de la renovación de la enseñanza médica en la Universidad de Valencia. Había sido, en 1721, uno de los firmantes de un memorial en el que se solicitaba autorización para enseñar la "medicina moderna" y era partidario del sistema iatromecánico, como se refleja en sus "aprobaciones" o prólogos a obras de José Arnau y del propio Piquer.

Recién graduado, Piquer participó en varias oposiciones a plazas del Hospital General de Valencia y a cátedras médicas de su Universidad. No tuvo éxito de momento, aunque la brillantez de sus ejercicios contribuyó a ir afianzando su prestigio profesional, que acrecentó la publicación de su primer libro, titulado *Medicina vetus et nova* (1735). En consecuencia, consiguió una excelente clientela y fue comisionado oficialmente para estudiar diversos problemas sanitarios. Los informes que redactó con este motivo reflejan la orientación que entonces tenía su práctica. Por ejemplo, aclaró que las muertes producidas a comienzos de 1739 en Fanzara eran debidas a "pulmonías o inflamaciones del pulmón", apoyándose principalmente "en una disección anatómica que se hizo (de un enfermo fallecido durante su estancia), en que se vio el pulmón inflamado en su lobo derecho donde sintió el dolor y circundado de una sania serosa sumamente acre que le cubría"

En 1741, Piquer entró en relación con Gregorio Mayáns Síscar, una de las figuras centrales de la Ilustración valenciana. La amistad entre ambos que entonces se inició se mantuvo, con algunos altibajos, hasta la muerte del primero. Mayáns influyó notablemente en Piquer, sobre todo en su formación humanística y filosófica, y también le ayudó de modo importante en su carrera. A la intervención de Mayáns se debió en parte que Piquer ganara en 1742 las oposiciones a la cátedra de anatomía de la Universidad de Valencia, poco después de resultar también vencedor en las de médico titular del Hospital General. El mismo año fundó Mayáns la Academia Valenciana, de la que Piquer fue uno de los primeros miembros. Esta institución no acabó de consolidarse y la participación del médico aragonés tampoco tuvo gran entusiasmo, expresando especialmente su disconformidad cuando Mayáns comenzó a imprimir las obras físicas y matemáticas de Juan Bautista Corachán que consideraba con razón anticuadas.

Durante los nueve años que ocupó la cátedra de anatomía valenciana, Piquer prestó gran atención a la enseñanza práctica, de acuerdo con la tradición de la escuela valenciana. "El año 1743 –afirmó en un texto suyo publicado dos años más tarde–, haciendo yo públicamente la demostración anatómica del hígado, una de las veinticinco disecciones prácticas que todos los años explico en cumplimiento de mi instituto, hallé en la cestilla de la hiel una piedra de tamaño de un huevo de gallina." Consta el mucho tiempo que le consumía la docencia y el prestigio que alcanzó ésta, lo cual no quiere decir que centrara su actividad científica en la morfología. Publicó algunos textos de tema anatómico y dejó otros inéditos, pero el objetivo fundamental de su obra escrita fue ofrecer "a la juventud española" una síntesis amplia "moderna" de los saberes médicos y de sus disciplinas básicas. A este programa responde la aparición del primer volumen de su *Física moderna, racional y experimental* (1745) y de su *Lógica moderna* (1747). Durante su etapa valenciana, Piquer publicó, además, un folleto (1745) en defensa de su *Física* y una serie de escritos polémicos (1746-48) en torno a un caso clínico que él había diagnosticado de "hectiquez", lo que implicaba la destrucción de los enseres que hubieran estado en contacto con el enfermo, de acuerdo con la legislación de la época relativa a la tuberculosis. Apareció entonces también la primera edición de su *Tratado de las calenturas* (1751), quizá su obra médica de mayor importancia.

A finales de 1751, Piquer marchó a Madrid como médico de cámara supernumerario de Fernando VI. Tradicionalmente se creía que en este nombramiento había influido Mayáns, pero Peset Llorca ha aclarado que Piquer fue a la corte reclamado por su maestro Antonio García, quien había sido nombrado médico regio y murió dos años después. El ascenso de Piquer fue muy rápido y en 1752 fue designado vicepresidente de la Real Academia Médico-Matritense (a la que pertenecía desde 1739) y entró a formar parte del Tribunal del Protomedicato. A lo largo de las dos décadas siguientes desplegó una notable actividad en ambas instituciones, de la que destacaremos los discursos que pronunció en la Academia sobre los sistemas médicos (1752, 1768, 1770), la tradición médica española (1761) y la historia de la medicina árabe (1770), así como sus informes desde el Protomedicato en torno a la inoculación de las viruelas (1757) y la reforma de la enseñanza médica (1767 y 1768).

Durante sus años juveniles, Piquer había sido seguidor de un eclecticismo fuertemente inclinado al sistema iatromecánico. En la primera edición de la *Medicina vetus et nova* (1735) afirma que la experiencia, la práctica y la razón le habían llevado a adherirse a las corrientes modernas y que, de todos los sistemas, estimaba certero "el que consideraba el cuerpo humano como una máquina constituída por sólidos y líquidos". En estos supuestos se basa la obra, que incluye un tratado de terapéutica y resúmenes sobre semiología de la orina y del pulso, y sobre piretología. Lo mismo sucede con sus *Theses medico-anatomicae* (1742), en las que concede gran importancia a la estequiología fibrilar, la indagación microscópica y la doctrina de la circulación de la sangre como núcleos del sistema iatromecánico. Continuó siendo partidario del mismo hasta el final de su etapa valenciana, tal como se refleja en la primera edición del *Tratado de calenturas* (1751), que lleva como subtítulo "según la observación y el mecanismo". Su

orientación personal estaba principalmente influida por Boerhaave y el cartesianismo médico, quedando en segundo término la obra de Baglivi y la tradición galileana.

El traslado a Madrid de Piquer coincidió con un cambio profundo de su mentalidad científica, porque en su discurso *De medicinae experimentalis praestantia* (1752) atacó ya de frente el iatromecanismo y defendió una postura antisistemática. La nueva actitud condicionó la modificación del contenido de la *Medicina vetus et nova* a partir de su tercera edición (1785) y del *Tratado de las calenturas* desde la segunda (1760), en la que incluso cambió de título. En ambas obras expuso los fundamentos de un "empirismo racional" opuesto a los sistemas cerrados y defensor ante todo de la importancia de la observación: "Es indubitable que la Medicina tuvo su principio por las observaciones, y lo es también que los progresos que ha hecho todos se deben a éstas... El raciocinio es el otro fundamento de la verdadera medicina, y para ser bien fundado ha de establecerse sobre buenas observaciones". Con este cambio, Piquer no hizo más que seguir una de las tendencias dominantes de la medicina europea de la época. Las obras que desde entonces citó más a menudo fueron las de las cabezas de la mentalidad antisistemática, en especial Gerhard van Swieten y otros miembros de la "Alte Wiener Schule", Giovanni Battista Morgagni, Albrecht von Haller, etcétera, así como las de Thomas Sydenham y los escritos hipocráticos estimados como modelos clásicos de observaciones clínicas. Por ello, publicó una excelente traducción de *Las obras de Hipócrates más selectas* (1757-1770), que comentó con "las observaciones prácticas de los antiguos y modernos" y destinó, conforme a su costumbre, a la "juventud española que se dedica a la medicina". Suavizados los entusiasmos innovadores de su época juvenil, subrayó que junto a las contribuciones modernas había que aprovechar las de los médicos antiguos. Este eclecticismo se refleja de modo muy acusado en las *Institutiones medicae* (1762) y la *Praxis medica* (1764-66), que escribió "ad usum Scholae Valentinae". Por otra parte, en la primera de estas obras, su extremada oposición a los sistemas le llevó a considerar la doctrina de la circulación sanguínea "no como cosa demostrada, sino más bien como una de las opiniones llamadas probables en las escuelas". También expresó su desconfianza en las observaciones microscópicas desde una postura que Terrada estima no muy alejada del sensualismo de Bichat: "Las observaciones certeras acerca de los objetos físicos –afirma Piquer– deben hacerse a través de varios sentidos, de forma que lo que se percibe con la vista también debe verificarse con el tacto y los demás sentidos".

La obra médica de Piquer alcanzó cierta difusión europea. Entre las ediciones extranjeras de sus textos destacan las de la *Praxis*, publicada en latín en los Países Bajos y en Italia, y las del *Tratado de las calenturas*, traducido al francés e impreso en dos ocasiones.

Menor relieve histórico tiene la *Física* (1745) de Piquer, a pesar de ser uno de los primeros libros con ese título publicados en castellano. Su principal mérito, según Víctor Navarro, consiste en ofrecer un nivel de información bastante actualizado, sobre todo si se tiene en cuenta que estaba destinado a los estudiantes de medicina y el enorme atraso que todavía imperaba entonces en las

---

32 EL LIBRO PRIMERO DE LAS EPIDEMIAS

XIV.

Κατὰ δὲ θέρος ἤδη καὶ φθινόπωρον, πυρετοὶ πολλοί, καὶ ξυνεχέες, οὐ βίαιοι· μακρὰ δὲ νοσέουσιν, οὐ δὲ περὶ τὰ ἄλλα δυσφόρως διάγουσιν, ἐγένοντο. Κοιλίαι ταραχώδεες τοῖσι πλείστοισι πάνυ εὐφόρως, καὶ οὐδὲν ἄξιον λόγου προσέβλαπτον. Οὖρά τε τοῖσι πλείστοισιν, εὔχροα μὲν, καὶ καθαρὰ· λεπτὰ δὲ, καὶ μετὰ χρόνον, περὶ κρίσιν πεπαινόμενα. Βηχώδεες οὐ λίην, οὐδὲ τὰ βησσόμενα δυσκόλως· οὐδ' ἀπόσιτοι, ἀλλὰ καὶ διδόναι πάνυ ἐνεδέχετο.

Jam vero ad aestatem & autumnum febres multae, assiduae, neque violentae prehendebant, istaque diu laborantibus, non his, qui caetera moleste habebant, contigerunt. Alvi plurimis valde placide conturbatae sunt, nihilque effatu dignae noxae

XIV.

En el Estío, y en el Otoño hubo muchas calenturas, las quales eran continuas; pero no fuertes, ni violentas, y estaban los pacientes largo tiempo enfermos, aunque *sin grande molestia*. A muchos de estos se les descomponia el vientre, y *lo llevaban bien*, y no experimentaban por ello daño digno de consideracion. Las orinas eran tambien en muchísimos de buen color, y *puras*, aunque de poco cuerpo; y andando el tiempo, quando se acercaba la crisis, salian cocidas. Tenian tós, pero no mucha, ni los fa-

XIV. Las calenturas, que Hippócrates propone en este texto, son las que comunmente llamamos quotidianas, las quales están descritas en nuestro *Tratado de Calenturas*, y se observan freqüentemente ácia el fin del Otoño, y principios del Invierno, en unos años mas que en otros, y en los que abundan de humores crasos con mas freqüencia, que en los coléricos. La benignidad, que se observó en estas calenturas, debe atribuirse á la constitucion del tiempo; porque hay algunos años, en que las fiebres quotidianas, aunque sean largas, son benignas, y al fin vienen á ceder á la naturaleza, y al tiempo: otros años son maliciosas, y poco á poco consumen á la naturaleza, y son inobedientes á toda suerte de remedios. En

Página de la edición de Piquer de **Las obras de Hipócrates más selectas** *(1757-1770). Incluye el texto original griego, su traducción al latín y al castellano, y un comentario principalmente de tipo clínico*



LIBER TERTIUS.

DE MORBIS INFIMI VENTRIS.

CAPUT I.

De Cardialgia.



ARDIAS apud Graecos cor, itemque os ſuperius ventriculi ſignificat, ſeu quod magnus ſit hujus cum corde conſenſus, vel quod eo affecto damnum facilè cordi communicetur. Igitur dolor ſuperioris orificii ventriculi *Cardialgia*: ejusdem partis morsus aut erosio *cardiogmos*: calor autem vehemens ſtomachi *cauſon* à quibuſdam nominatur; ſed harum vocum diſtinctio nimium ſubtilis confuſionem potius quàm perſpicuitatem parit, praeſertim cum invaluerit uſus eas omnes affectiones ſub uno *Cardialgiae* nomine ſignificandi (a). Eſt ergo *Cardialgia*: *triſtis & moleſtus oris ſuperioris ventriculi ſenſus, nunc dolore, nunc morsu, perpetuò autem anxietate, viriumque impotentia conjunctus.* Diſcimus hinc, non omnem ventriculi dolorem cardialgiam eſſe, ſed eum dumtaxat, qui angoribus, mor-

A ſi-

(a) Vide Foeſium *Oeconom. Hipp. verb.* καρδιαλγια, *pag.* 305. Sennert. *Pract. lib.* 3. *part.* 1. *ſec.* 2. *cap.* 15. & 16. *pag.* 408. *ſeq.*

Página de las **Institutiones Medicae** *(1762), de Andrés Piquer. Resulta patente la belleza de su impresión por Joaquín Ibarra, el mejor tipógrafo de la España del siglo XVIII*

universidades españolas en estas materias. Piquer sigue un criterio ecléctico, citando principalmente a Descartes, Pierre Gassendi, Robert Boyle y también Newton, cuyas teorías conocía indirectamente a través de sus divulgadores franceses. En conjunto, el libro adolece de un excesivo empirismo y de incomprensión del papel de la matemática como lenguaje insustituible de la física moderna. Por otro lado, su descalificación del sistema de Copérnico excede con mucho las moderadas cautelas de José de Zaragoza y Tomás Vicente Tosca.

De la *Lógica* (1747) de Piquer solamente anotaremos que es una mezcla de lógica clásica y teoría moderna del conocimiento. Sus fuentes son fundamentalmente las obras de Aristóteles, de Descartes y otros racionalistas y, sobre todo, de Locke y los empiristas ingleses. Según Blasco, es un intento asistemático y poco desarrollado, pero interesante, de poner la lógica al servicio de la ciencia.

Esperanza de vida y crecimiento demográfico en el Tercer Mundo

El rápido descenso de la tasa de mortalidad tras la Segunda Guerra Mundial produjo un crecimiento explosivo de la población. El ritmo ha remitido. Mas aunque reanudara su velocidad apenas afectaría ya al tamaño de la población

Davidson R. Gwatkin y Sarah K. Brandel

¿Deberían los países industrializados hacer un enérgico esfuerzo por ayudar a reducir el índice de mortalidad en el Tercer Mundo? La esperanza de vida de los países en vías de desarrollo se cifra en 15 o 20 años menos que en Europa y en Estados Unidos. Desde el punto de vista humanitario, sería justo que las diferencias se acortaran sin demora. Téngase presente, sin embargo, que el rápido aumento de la esperanza de vida registrado en los países en vías de desarrollo desde la Segunda Guerra Mundial ha sido una de las principales causas del crecimiento demográfico. Y semejante crecimiento, espectacular, agrava seriamente los problemas económicos y sociales. Es, pues, razonable preguntarse si un incremento sustancial y rápido de la esperanza de vida en el Tercer Mundo no empeoraría las cosas.

Las mejores estimaciones en boga indican que la población del Tercer Mundo se triplicará, a buen seguro, en los próximos cien años; a partir de cuyo momento permanecerá estable. Aun cuando la esperanza de vida creciera al ritmo que se presenta como el más rápido posible, el efecto final sobre la población estable del Tercer Mundo sería casi inapreciable. La razón de ello estriba en que el índice de crecimiento de los países en vías de desarrollo es cada vez menos sensible a las variaciones del índice de mortalidad. Los factores que más pueden influir sobre el crecimiento demográfico son las tendencias futuras de la fecundidad y el gran número de jóvenes que actualmente alcanzan la madurez sexual, principalmente a causa de la alta fecundidad del pasado reciente. Si hay que mantener a un mínimo el crecimiento de la población, importará más tratar de reducir el índice de natalidad. Ninguna razón de índole demográfica tiene el peso suficiente para impedir que los países industrializados responsables ayuden a aumentar la esperanza de vida de los países en vías de desarrollo.

El rápido crecimiento de la población en los países en vías de desarrollo de Africa, Asia e Iberoamérica es consecuencia del descenso de la mortalidad que marca la primera fase de la transición demográfica. La transición se considera generalmente el acontecimiento fundamental de la historia reciente de la población humana. Durante la transición, una población evoluciona de un equilibrio de alta fecundidad y alta mortalidad a un equilibrio de baja fecundidad y baja mortalidad.

Antes de la transición demográfica, el índice de mortalidad era lo suficientemente alto para evitar un crecimiento significativo a largo plazo, a pesar de la alta fecundidad. Esta situación ha prevalecido durante la mayor parte de la historia humana. En Europa occidental, el índice de mortalidad comenzó su descenso en los siglos XVII y XVIII. En muchos países, la fecundidad se mantuvo alta hasta varias décadas después de iniciado el descenso de la mortalidad, lo que dio lugar a un rápido crecimiento de la población. El incremento anual en muchos países se situó entre el 0,5 y el 1 por ciento. Finalmente, el índice de mortalidad alcanzó un nivel tan bajo que difícilmente podían esperarse mayores progresos. El índice de natalidad empezó a disminuir durante la transición de la sociedad agrícola a la sociedad industrial. El crecimiento de la población se frenó cuando los índices de natalidad y mortalidad llegaron a igualarse en un mismo nivel bajo.

Una medida ampliamente utilizada para calcular la mortalidad es la esperanza de vida al nacer. Representa el número de años que se espera viva un niño según los niveles de mortalidad que prevalezcan en el momento de su nacimiento. La esperanza de vida al nacer en la mayoría de los países europeos está actualmente entre los 72 y los 73 años, más del doble de los valores anteriores al comienzo de la transición demográfica. Una medida correspondiente de los nacimientos es el índice de fecundidad total, que viene dado por el número de hijos que tendría una mujer a lo largo de toda su vida fecunda si se atuviese al número de descendientes que señala el índice de natalidad para cada edad. En Europa, a finales del siglo XVIII, el índice medio de fecundidad total era probablemente de cinco hijos; hoy se sitúa entre 1,4 y 2,6, aproximadamente. Puesto que no todos los niños alcanzarán la edad de reproducción, para mantener la población constante en una sociedad con un nivel de mortalidad europeo se requiere un promedio de cerca de 2,1 hijos por mujer. En gran parte de Europa, pues, la fertilidad está en el nivel de reposición, o por debajo de él.

Una transición similar hacia mortalidades y fecundidades bajas se ha producido en países industrializados no europeos: Australia, Canadá, Japón y Estados Unidos. Se está dando ahora en Argentina, Cuba, Hong Kong y Singapur, donde el desarrollo económico ha alcanzado un nivel bastante alto. En la mayoría de los países en vías de desarrollo de Africa, Asia e Iberoamérica, la transición demográfica, a pesar de todo, no ha llegado muy lejos. El índice de mortalidad ha bajado rápidamente en los últimos 25 años o más, pero el descenso de la fecundidad acaba de comenzar. Esa disparidad en los índices es la causa del acelerado crecimiento actual de la población. Queda por ver si la evolución de la transición demográfica en el Tercer Mundo se asemejará a la que tuvo lugar en Europa.

Desde la Segunda Guerra Mundial,

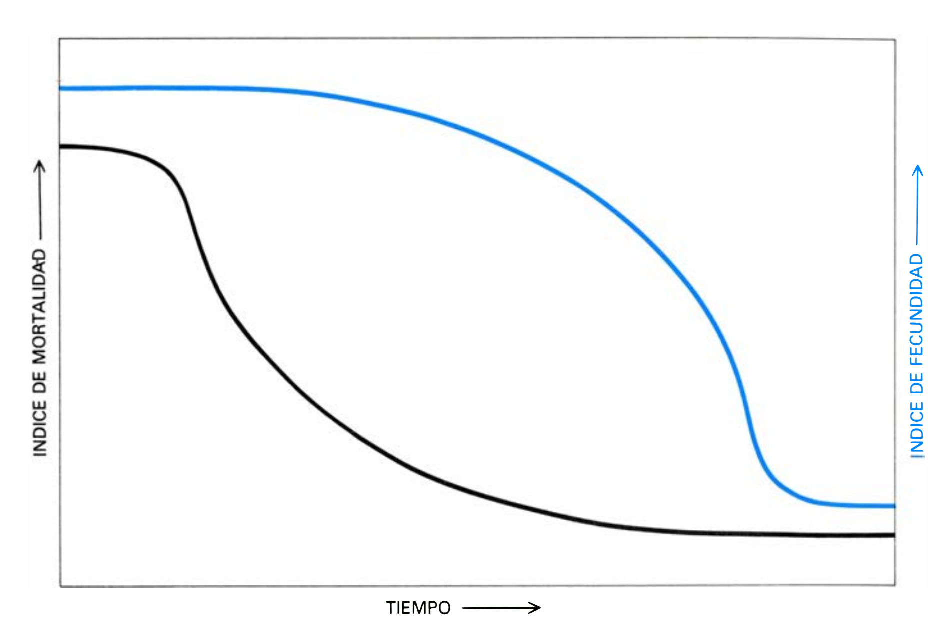
el progresivo descenso del índice de mortalidad ha llevado a un incremento de un 50 por ciento de la esperanza de vida en los países en vías de desarrollo: de menos de 40 años a 55 años, aproximadamente. Hasta mediados de los sesenta, la fecundidad cambió poco; comenzó luego a descender en gran parte del Tercer Mundo, excepto en el Africa subsahariana. A finales de los setenta, el índice de fecundidad total medio en los países en vías de desarrollo rondaba el 4,7, un 20 por ciento menos que 15 años antes.

Sin embargo, ese descenso de la fecundidad no basta para compensar el efecto de las grandes reducciones de mortalidad de los años anteriores. El índice de fecundidad total en el Tercer Mundo sigue siendo aproximadamente el doble que el índice de reposición. Además, los descensos de fecundidad tardan algún tiempo en afectar al crecimiento demográfico; de ahí que la población del Tercer Mundo haya persistido en una línea de rápido ascenso y prosiga en esa tendencia hasta bien entrado el siglo próximo, aun cuando la fecundidad continúe en su curva descendente. El índice de crecimiento anual es, en la actualidad, algo superior al 2 por ciento, lo que duplica la población cada 33 años. En algunas regiones ese índice alcanza valores que duplican la población en menos de 20 años.

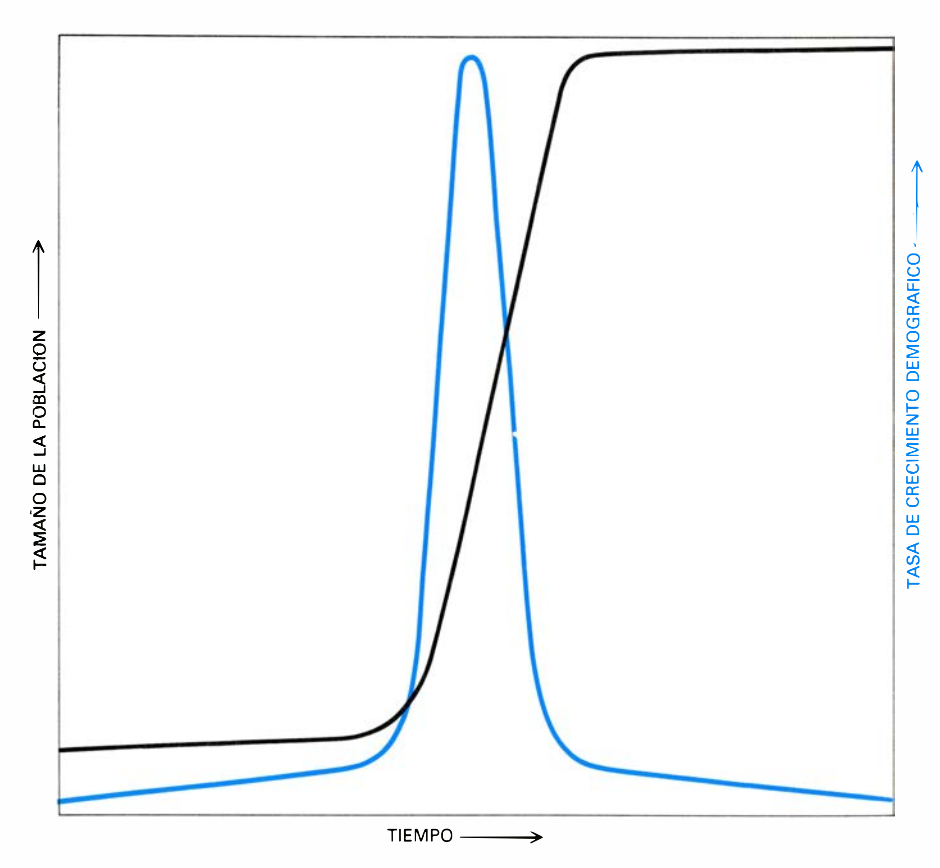
El índice de crecimiento que hoy se da en los países en vías de desarrollo es varias veces superior al que en una fase comparable de su transición demográfica tenían los países europeos. Se explica en parte esa disparidad por el rápido descenso de la mortalidad registrado en el Tercer Mundo. La esperanza de vida se ha incrementado entre 2 y 5 veces más deprisa que en la Europa de los primeros estadios de la transición.

En los seis países de Europa occidental de los que se dispone de datos adecuados, la media de la esperanza de vida creció de los 40 a los 50 años entre 1830 y 1900, aproximadamente. El posterior progreso habido en este siglo la situó en unos 65 años hacia la década de los cincuenta. Desde entonces se ha incrementado más lentamente, hasta llegar al valor actual de 73 años.

En la Europa meridional y oriental, el incremento de la esperanza de vida empezó más tarde; poco antes de 1900 se cifraba todavía en torno a los 40 años. Sin embargo, una vez iniciado, el incremento fue más rápido que el que se produjo en Europa occidental. El paso de los 40 a los 50 años se dio hacia el primer cuarto del siglo XX. Siguieron luego incrementos más rápidos incluso,



SE SUELE CONSIDERAR LA TRANSICION DEMOGRAFICA como el principal acontecimiento de cuantos configuran la historia reciente de la población humana. Prácticamente concluida en los países industrializados, se encuentra a mitad de camino en los países en vías de desarrollo, donde su resultado es todavía incierto. Se indica esquemáticamente el efecto de la transición sobre la mortalidad y la fecundidad. Antes de la transición se hallan aproximadamente en equilibrio un alto índice de natalidad con un alto índice de mortalidad. Al concluir la transición, un bajo índice de natalidad queda compensado con un bajo índice de mortalidad. El descenso de la mortalidad precede generalmente al descenso de la fecundidad; durante el intervalo, la población puede crecer rápidamente. El descenso de la mortalidad en los países en vías de desarrollo se acentuó a partir de la Segunda Guerra Mundial. El descenso de la fecundidad no empezó a generalizarse hasta los años 60; desde entonces la fecundidad se ha reducido en un 20 %.



TAMAÑO DE LA POBLACION E INDICE DE CRECIMIENTO muestran la influencia de la transición demográfica en esta representación esquemática. Antes de la transición, la población es pequeña y estable. Cuando baja la mortalidad, sube la tasa de crecimiento. La población continúa creciendo rápidamente durante un período considerable hasta que la reducción de la fecundidad logra frenar el ritmo de crecimiento. Algún tiempo después de que la mujer media tenga 2,1 niños, aproximadamente, la población se estabiliza otra vez con un tamaño mucho mayor. En el siglo XIX europeo, el índice anual de crecimiento demográfico raramente excedía del 1 por ciento. En el Tercer Mundo, donde la fecundidad era muy alta cuando la mortalidad empezó a bajar, se han registrado índices de crecimiento anual bastante altos, hasta de un 4 por ciento. La tasa global es ahora de un 2,1 por ciento. Algunas de las previsiones en boga sugieren que cuando se alcance la estabilidad, la población en el conjunto de los países en vías de desarrollo, que actualmente suma unos 3000 millones, será, aproximadamente, de unos 9000 millones.

anulándose pronto la diferencia entre la Europa suroriental y la occidental. Hacia los años 1970, la esperanza de vida en el sudeste europeo era casi la misma que en poniente.

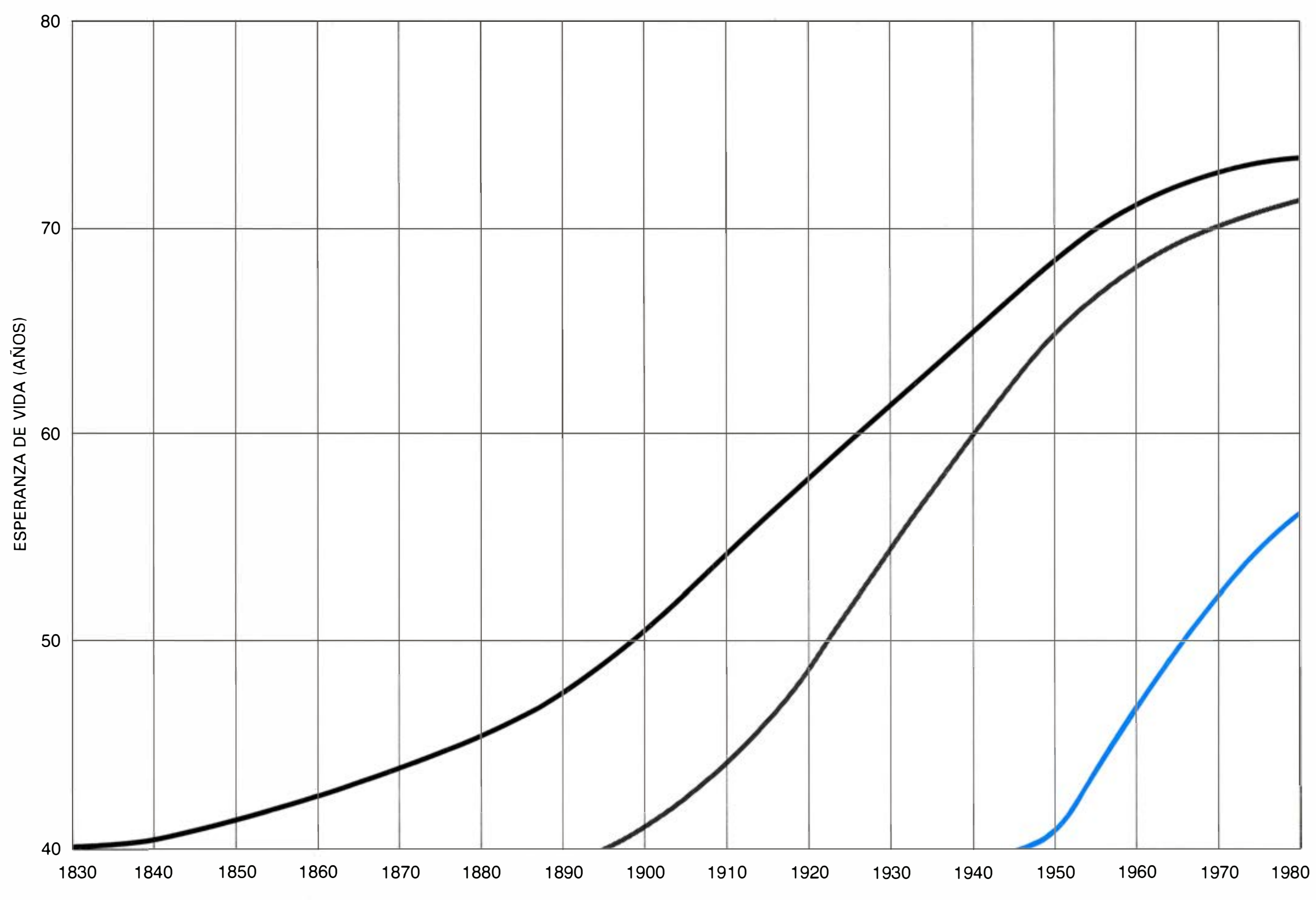
La mortalidad no inició su curva descendente en los países en vías de desarrollo de Africa, Asia e Iberoamérica hasta que el proceso estaba ya bastante avanzado en los países industrializados. Aunque en algunos en vías de desarrollo, así la India y varias naciones sudamericanas, las reducciones empezaron a principios del siglo xx, en la mayoría de los países subdesarrollados no comenzó una disminución sustancial hasta pasada la Segunda Guerra Mundial.

Hacia 1950, la media de esperanza de vida de un niño nacido en el Tercer Mundo giraba en torno a los 40 años. A partir de entonces, y durante una década, se observó un crecimiento anual de la esperanza de vida equivalente a un año o más en países tan diferentes como Jamaica, Malasia, Mauritania, México, Sri Lanka y Taiwan. En China continental se mantuvo durante tres décadas un incremento similar. La esperanza de vida en el conjunto de los países en vías de desarrollo alcanzó el medio siglo hacia 1965. En 15 años, el Tercer Mundo había recorrido un camino que a Europa le costó 70.

La disminución del índice de mortalidad en los países en vías de desarrollo, aunque rápida, no fue uniforme: fracasó en el intento de reducir las diferencias entre regiones. Según informes de las Naciones Unidas, en 1950 Iberoamérica tenía la más alta esperanza de vida de las tres principales regiones en vías de desarrollo: unos 52 años. Su incremento anual durante los años cincuenta se situó entre 0,5 y 0,6 años. En Asia, la esperanza de vida en 1950 rondaba los 43 años, pero el incremento anual subsiguiente fue mayor: unos 0,7 años. Africa tenía la esperanza de vida inicial más baja, 37 años, y el subsiguiente incremento anual fue también el más pequeño, entre 0,4 y 0,5 años. A mediados de los setenta este aumento situó la esperanza de vida del continente africano en unos 47 años; en Asia era de 57 y, de 62, en Iberoamérica.

El incremento anual medio de la esperanza de vida en todos los países en vías de desarrollo, durante los años cincuenta, fue de más de 0,6 años. Sin embargo, en la década siguiente, no se mantuvo ya ese ritmo. Según estimaciones de las Naciones Unidas, hacia finales de los sesenta y principios de los setenta el aumento anual bajó aproximadamente a 0,4 años. La deceleración del ritmo con que subía la esperanza de vida afectó a las tres regiones. Ese freno impide cerrar, a la velocidad de antes, la brecha que separa a los países industrializados de los que están en vías de desarrollo, brecha cifrada hoy en unos 17 años.

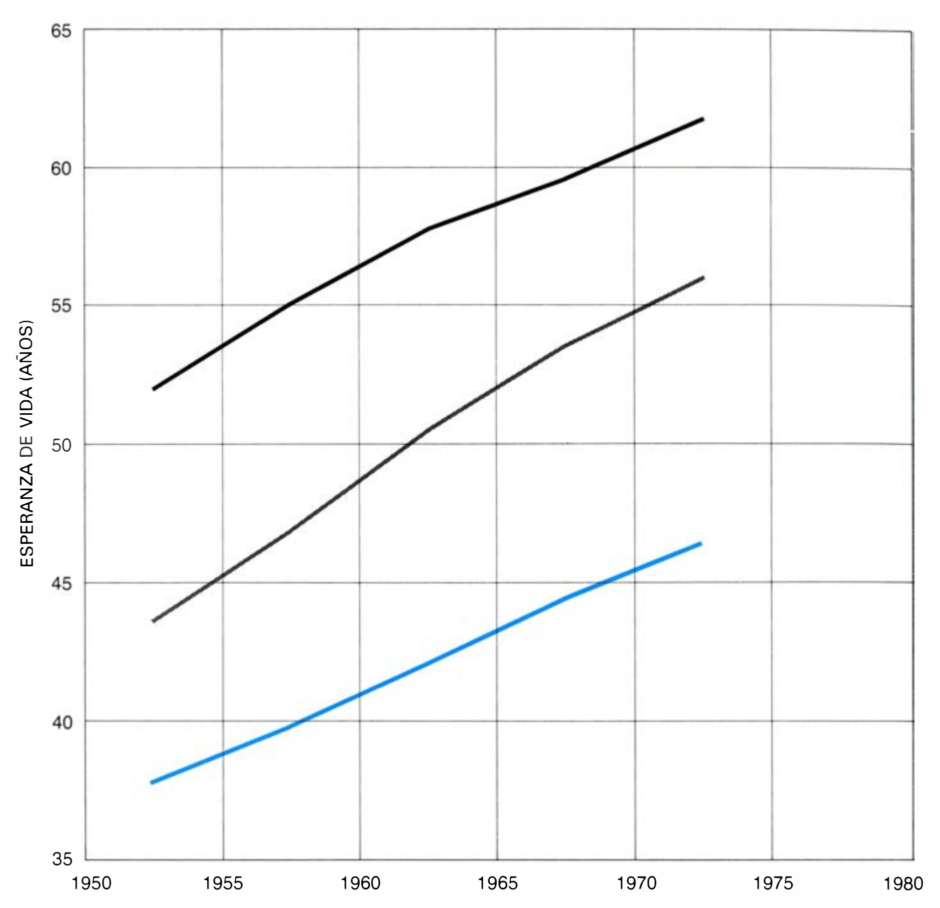
No se conocen con certeza las razones de ese enlentecimiento del aumento de la esperanza de vida; factores sanitarios, sociales y económicos parecen haber intervenido de una manera destacada. Las medidas sanitarias tuvieron sólo una influencia limitada en la evolución de la esperanza de vida en la Europa del XIX. La mejora de la alimentación y de las condiciones de vida, resultante de una mayor productividad



UN AUMENTO DE LA ESPERANZA DE VIDA de 40 a 50 años tiene lugar en los primeros estadios de la transición demográfica. Ese aumento de la longevidad media de la población se alcanzó en Europa occidental (*curva negra*) entre 1830 y 1900, aproximadamente. Una segunda ola de baja de mortalidad en el Sur y el Este de Europa (*curva gris*) empezó más tarde, pero fue también más veloz; el aumento de 40 a 50 años tuvo lugar en los primeros 25 años del siglo xx. Hacia la década de 1970, la esperanza de vida era de más de 70 años en casi toda Europa. La tercera ola de freno de la mortalidad, que corresponde a los países en vías de desarrollo (*curva de color*), fue la más tardía, aunque la más rápida de todas en sus primeras fases. En 1950, la esperanza de vida en los países en vías de desarrollo era de unos 40 años. Hacia 1965 subió a 50 años, consiguiendo en 15 años lo que había requerido unos 70 en Europa occidental. La esperanza de vida en el Tercer Mundo se sitúa hoy en torno a los 55 años, de 15 a 20 años menos que en los países desarrollados.

agrícola y desarrollo económico, parece haber sido la principal causa del descenso de la mortalidad. En cambio, la inmediata introducción en los países en vías de desarrollo de medicamentos e insecticidas descubiertos durante la Segunda Guerra Mundial, que permitió controlar muchas enfermedades contagiosas, parece ser la causa de un 50 por ciento del incremento de la esperanza de vida en el Tercer Mundo. Esas medidas elevaron la esperanza de vida por encima del nivel que se hubiera alcanzado a través del progreso social y económico.

Pero se diría que los medicamentos y pesticidas hubieran empezado a agotar sus efectos. Las enfermedades contra las que actúan están casi todas bajo control. Quizá sea esa una de las razones de la disminución del aumento de la esperanza de vida. Se ha comprobado que las enfermedades que subsisten son más difíciles de controlar. Quizá alguna de las numerosas y prometedoras tentativas que se están poniendo en práctica produzca algún cambio. Una posibilidad consiste en preparar personal no especializado para administrar medicamentos básicos en áreas rurales. En otra se enseña a las madres de niños afectados de diarrea a administrarles una solución líquida barata para combatir la deshidratación. También se trabaja en el desarrollo de vacunas contra enfermedades parasitarias.



LAS REGIONES DEL TERCER MUNDO muestran importantes diferencias en la esperanza de vida. A principios de los años 50, el africano medio (*color*) podía esperar vivir unos 37 años, el asiático medio (*gris*) unos 43, y, el iberoamericano medio (*negro*), unos 52. El período comprendido entre 1950 y 1975 registró un rápido aumento de la esperanza de vida en esos tres puntos. Dado que la magnitud del incremento fue similar en los tres, se mantuvieron las disparidades. A principios de los 70, la esperanza de vida en Africa era de unos 47 años, en Asia, de unos 57 años, y, en Latinoamérica, de unos 62.

La diarrea, la neumonía y la desnutrición, que son las causas principales de mortalidad entre los jóvenes de los países en vías de desarrollo, no pueden solucionarse únicamente con medidas sanitarias. Están asociadas a un bajo nivel de vida; una mejora en las dietas, un buen suministro de agua, un hábitat digno, ropas y equipamientos sanitarios son imprescindibles para controlarlas. La aparente retención del ritmo de desarrollo en muchos lugares del Tercer Mundo, después del progreso rápido de los años 50 y 60, se halla también en la base de ese amortiguamiento registrado en la curva ascendente de la esperanza de vida.

No es fácil avivar el ritmo de desarrollo, pero los datos disponibles sugieren que el aumento potencial de la esperanza de vida podría ser mayor. Los países con una renta per cápita anual de 450 dólares, por ejemplo, parecen tener una esperanza de vida de unos 16 años más que los países con una renta per cápita de 220 $. A igualdad de condiciones, un país pobre con una distribución equitativa de la renta ofrece una esperanza de vida de unos 6 años más que otro en el que la renta esté mal distribuida. La educación, particularmente de la mujer, constituye un factor importante. Aunque incrementar el ritmo de aumento de la esperanza de vida en el Tercer Mundo pueda requerir un enorme esfuerzo, no hay razones para pensar que sólo puedan conseguirse incrementos lentos. Quizá no se sobrepase el rápido avance de los años cincuenta, pero el progreso social y económico y unas medidas sanitarias producirían incrementos más rápidos que los del pasado reciente.

Los fuertes incrementos anuales de la longevidad alcanzados en los años 50 fueron una de las causas principales del rápido aumento de la población; conviene, pues, estudiar la relación que guardan los incrementos de la esperanza de vida y el crecimiento de la población. Aun cuando el aumento de la esperanza de vida se acelera de nuevo, es improbable que ello provoque un crecimiento de la población tan rápido como el registrado con los aumentos anteriores. Por dos razones.

La primera es que las reducciones de la mortalidad tienen una influencia menor sobre el crecimiento de la población a medida que se alcanzan más altos niveles de esperanza de vida. La reducción del efecto se debe al cambio en la distribución de la edad de fallecimiento. Un niño salvado de una viruela letal puede vivir 50 o 60 años y morir de alguna otra causa. A una madre que hubiera muerto de parto le restarán todavía 30 o 40 años de vida. Una persona de 70 años que padezca de insuficiencia coronaria, dispondrá de cinco años más de vida. Como consecuencia de este retraso de la muerte, sube la media de edad de los fallecidos, junto con la proporción de muertes entre los más viejos.

La esperanza de vida no indica en absoluto a qué edad muere la mayoría de la gente. Cuando la esperanza de vida es de 35, por ejemplo, muy poca gente vive hasta los 35 y fallece al día siguiente. Al nivel de fecundidad que ahora predomina en el Tercer Mundo, del 35 al 40 por ciento de todas las muertes son de niños de 5 años o menos; del 25 al 30 son de sexagenarios o mayores. Al incrementarse la esperanza de vida, el cambio causado por el retraso de la mortalidad produce una modificación constante de este modelo. Cuando la esperanza de vida llega a 75, al nivel de

fecundidad actual, sólo de un 5 a un 20 por ciento de las muertes son de niños de 5 años o menos y de un 70 a un 80 son de personas de 60 años o más.

Este cambio comporta consecuencias significativas para el crecimiento de la población, porque raramente los ancianos tienen hijos. Por ello, retrasar la muerte de un sexagenario añade sólo una persona a la población. Un niño salvado representa el principio de una cadena. En años posteriores, sus hijos, nietos y futuros descendientes se añadirán a la población en cantidades que aumentan geométricamente. La gente de mediana edad conserva aún cierta capacidad de reproducción, de ahí que el efecto de retardar la muerte de una persona de mediana edad sea mayor que el de salvar un anciano, pero menor que el de salvar un niño.

Cuando la esperanza de vida es baja y una gran proporción de los que mueren son jóvenes, un incremento de la duración de la vida significa generalmente que gran parte de los que se salvan son jóvenes. El efecto sobre el crecimiento es importante, pues los jóvenes tienen aún por delante su período de reproducción. No mueren muchos jóvenes cuando se alcanzan niveles más altos de esperanza de vida. Si la esperanza de vida se incrementa todavía más, el aumento se conseguirá alargando la vida de los mayores, que probablemente no tengan ya más hijos. En este caso, pues, el crecimiento de la población será mínimo.

Por todo ello, a medida que aumenta la edad de las personas "rescatadas", los incrementos de la esperanza de vida tienen un efecto progresivamente menor sobre la población. Un aumento de 5 años a partir de los 55 (el nivel medio actual en el Tercer Mundo), conllevaría un incremento de la población de sólo un poco más de la mitad que un aumento de 5 años con una esperanza de vida de 35. Como han demostrado Ansley J. Coale, de la Universidad de Princeton, y posteriormente Nathan Keyfitz, de Harvard, un incremento de la esperanza de vida desde los 70 hasta más de 100 años tendría un efecto insignificante sobre el crecimiento de la población a largo plazo.

La segunda razón por la cual no cabe esperar que un incremento más rápido de la esperanza de vida afecte sensiblemente al tamaño de la población es la disminución de la fecundidad, que ha empezado ya en el Tercer Mundo y que se espera que continúe. En los años 50 y 60, cuando la mortalidad decreció al máximo, no se dio un cambio importante en los altos niveles de fecundidad de los países en vías de desarrollo. Sin embargo, desde mediados de los 60, la fecundidad de la mayoría de los países en vías de desarrollo ha descendido un tanto. La mayoría de demógrafos espera que, hacia finales de siglo, el descenso comporte reducciones adicionales de un 20 a un 35 por ciento en la totalidad del Tercer Mundo.

El descenso de la fecundidad ejerce, obviamente, un efecto directo sobre el crecimiento de la población. Como hemos indicado más arriba, los niños supervivientes no sólo afectan al tamaño



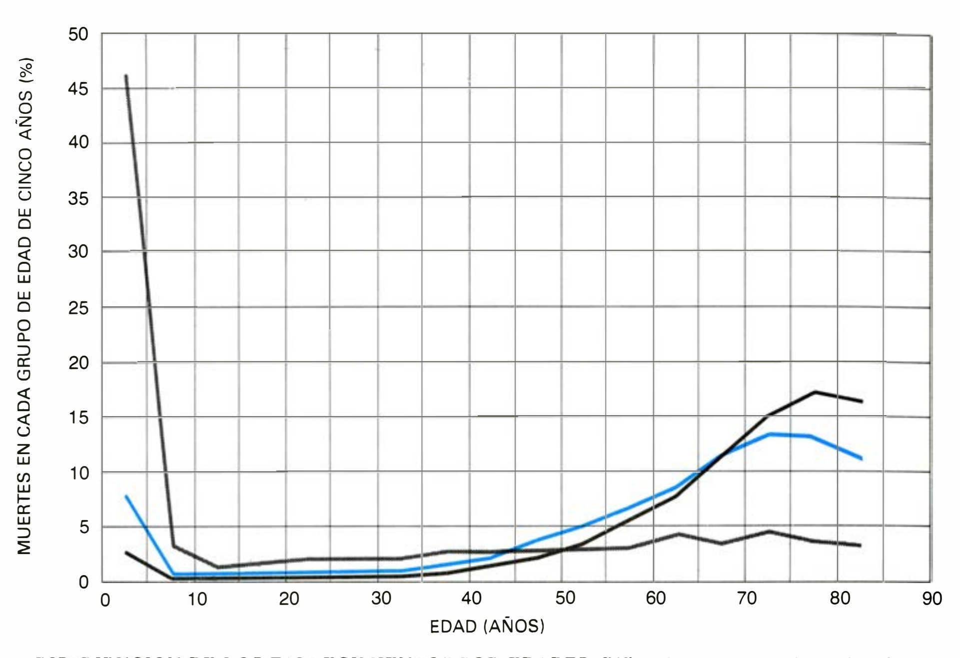
DISMINUCION DEL RITMO de reducción de la mortalidad en años recientes en los países en vías de desarrollo. Queda reflejada en la media del incremento anual de la esperanza de vida en las grandes regiones. Los datos están tomados de las estimaciones efectuadas por las Naciones Unidas. Durante los años 50, la esperanza de vida aumentó anualmente más de 0,6 años en el Tercer Mundo. Este índice no tenía precedentes. Hacia 1975, el incremento anual había bajado a 0,4 años. Se cree que la disminución se debe a dos factores. Uno, que las medicinas y pesticidas introducidos en los países en vías de desarrollo después de la Segunda Guerra Mundial permitieron controlar un gran número de las enfermedades contra las que son eficaces. Las enfermedades que subsisten son más difíciles de erradicar con los métodos al alcance. La segunda razón es que el ritmo de desarrollo social y económico, que guarda una relación directa con el descenso de la mortalidad, ha disminuido en muchos países en vías de desarrollo.

de la población por sí mismos, sino porque, a su vez, tienen descendientes. A medida que disminuye la fecundidad nacen menos niños y cada niño tiene, por término medio, menos descendientes, de ahí que la supervivencia de esos niños tenga cada vez menores consecuencias para el crecimiento. Además, el descenso de la fecundidad contribuye a un cambio ascendente en la media de edad de fallecimiento. Si nacen menos niños, morirán menos, de donde una mayor proporción de muertes corresponderá a personas ancianas.

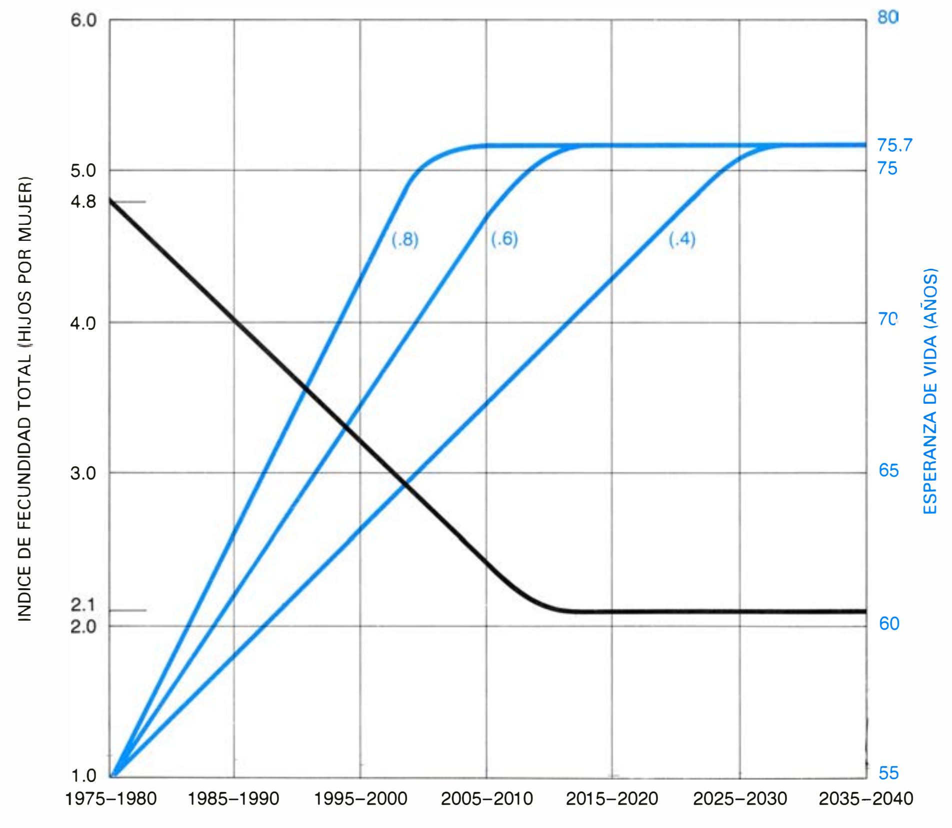
Hay, pues, dos argumentos que sugieren que el efecto de una reducción más rápida en el índice de mortalidad sería menor que el efecto del rápido descenso anterior. Para determinar el efecto es necesario simular matemáticamente la incidencia, sobre el tamaño de la población, de varios índices de cambio de la mortalidad. Conviene, por comparación, simular también el efecto de los cambios de la fecundidad. Por ello se requiere información, o la formulación de supuestos, sobre los índices de cambio de la fecundidad y la mortalidad, la distribución por edades de los descensos de mortalidad y fecundidad y el tamaño inicial de población.

Valen como puntos de partida los supuestos sobre las futuras tendencias de la mortalidad y fecundidad efectuados por demográfos de las Naciones Unidas, la Oficina norteamericana del Censo y el Banco Mundial. Las estimaciones más utilizadas probablemente sean las de las Naciones Unidas. La proyección media de 1978 de las Naciones Unidas para la población mundial se basa en el supuesto de que el índice total de fecundidad disminuirá de 4,7 a 3,4 y la esperanza de vida aumentará de 55 a 63 años hacia finales de siglo. Ello equivale a una disminución anual del índice de fecundidad total de un 0,07, aproximadamente, y a un incremento anual de unos 0,4 años en la esperanza de vida.

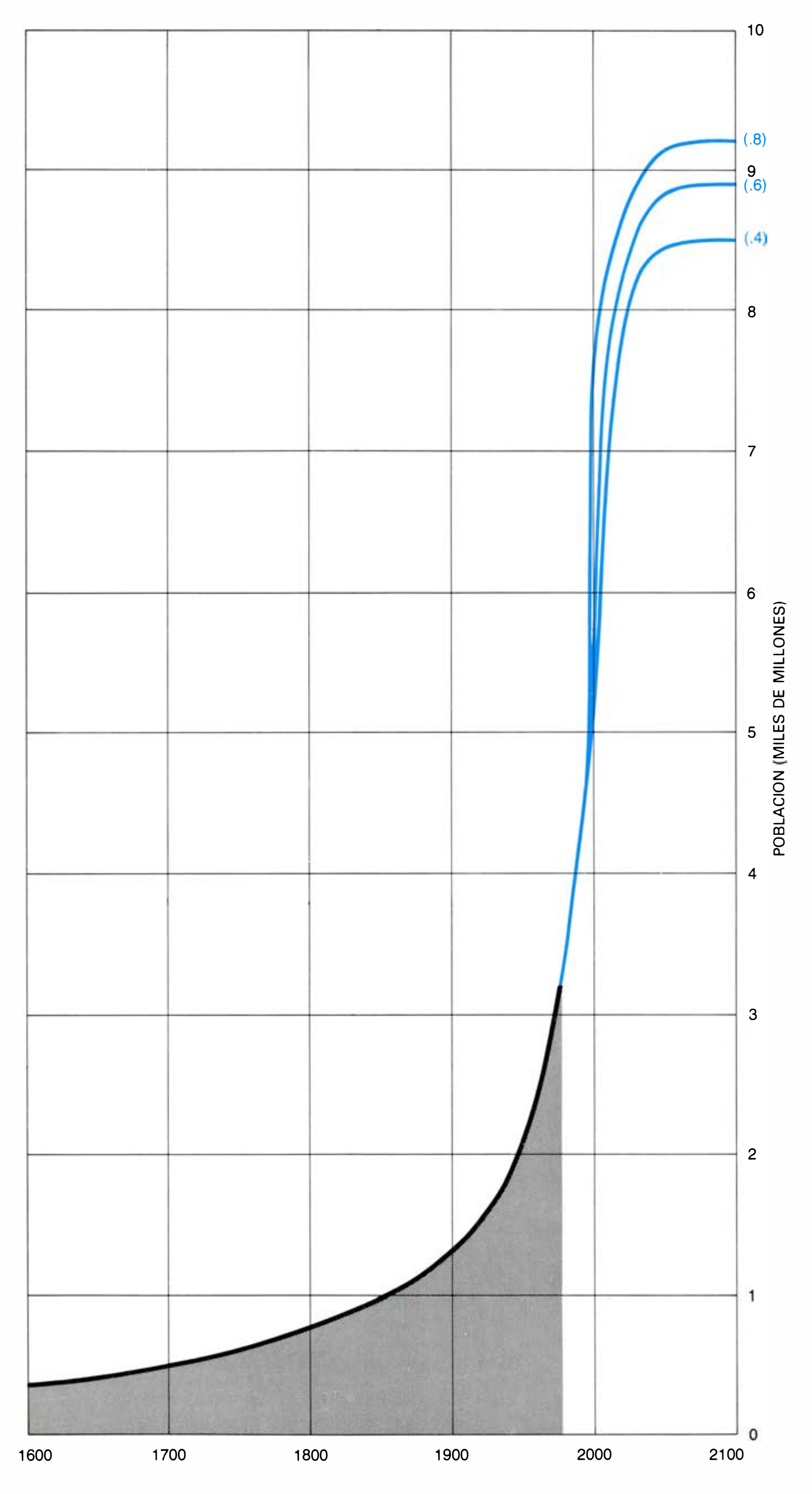
De los trabajos de Coale y Paul Demeny, del Consejo de Población, que suelen servir de referencia, se deduce una distribución por edades del descenso de la mortalidad. Coale y Demeny elaboraron tablas de vida ideales a partir de muchas fuentes de datos sobre la mortalidad en Europa y en otros lugares. Las tablas de vida se agrupan en cuatro familias, que representan distintos modelos de distribución de fallecimientos por edad y sexo a varios niveles de esperanza de vida. Utilizamos la familia "occidental", que se basa en datos de Australia, Canadá, Europa occidental, Israel, Japón, Nueva Zelanda,



DISTRIBUCION DE LOS FALLECIMIENTOS POR EDADES. Difiere bastante entre los países desarrollados y los países en vías de desarrollo. Las curvas muestran la proporción de muertes en cada grupo de edad de 5 años en México (*gris*), Suecia (*negro*) y Estados Unidos (*color*) en 1960. En las naciones en vías de desarrollo, ejemplificadas por México, predominan las muertes infantiles. En 1960, más del 45 por ciento de todas las muertes fueron de niños menores de 5 años. En el curso del desarrollo económico, el índice de natalidad y el índice de mortalidad bajan y la estructura de edad de la población cambia. Predomina la población anciana y, por tanto, el mayor número de muertes se da en este colectivo. Salvar la vida de los niños tiene un efecto mucho mayor sobre el crecimiento de la población que alargar la vida de los viejos. Aumentar la esperanza de vida de 35 a 55 años (cosa que ya ha ocurrido en el Tercer Mundo) provoca un crecimiento demográfico mucho mayor que el hecho de aumentarla de 55 a 75 (lo que resta aún por hacer). El efecto se refuerza con el descenso de la fecundidad, que reduce la mortalidad infantil.



SUPUESTOS FORMULADOS POR LOS AUTORES al proyectar el tamaño de la población futura del Tercer Mundo. El número de niños que dé a luz la mujer media descenderá presumiblemente de forma líneal a partir del nivel inicial de 4,8 hasta el nivel de reposición que se registre hacia el año 2020. Tres índices anuales diferentes de aumento de la esperanza de vida se emplearon al objeto de poder contrastar sus distintos efectos sobre el crecimiento de la población. El primer índice mantiene el aumento anual actual de 0,4 años, alcanzando el máximo de esperanza de vida de unos 76 años hacia el 2025. El segundo índice anual es de 0,6 años, una vez y media el índice actual y justo por debajo de la elevada tasa de los años 50; a este ritmo, el máximo de esperanza de vida se alcanzaría en el año 2010. El tercer índice, 0,8 años, es decir, el doble del índice actual, nos lleva a la máxima longevidad un poco después del año 2000.



LA POBLACION ESTABILIZADA que se alcance en el Tercer Mundo será mucho mayor que la población actual, pero variará poco con el índice de aumento de la esperanza de vida. El gráfico muestra las proyecciones de los autores en torno a la población futura, basadas en tres índices de aumento de la longevidad. La parte sombreada muestra los datos históricos. La población del Tercer Mundo ronda los 3200 millones. De acuerdo con las proyecciones se estabilizará hacia el 2100. Si el actual índice anual de aumento de la esperanza de vida (0,4 años) persiste, el tamaño de población estabilizada será de 8500 millones. Si el índice aumenta a 0,6 años, la población será de 8900 millones. En el caso de que aumente a 0,8 años, la población será de 9200 millones. La tasa de 0,6 años es probablemente la más elevada que pueda alcanzarse en las condiciones actuales. Nos lleva a una población final sólo un 6 por ciento mayor que la que se prevé en el índice de la línea de base. Los autores indican que la diferencia real podría ser menor, porque un descenso de la mortalidad conduciría a reducciones más rápidas de la fecundidad.

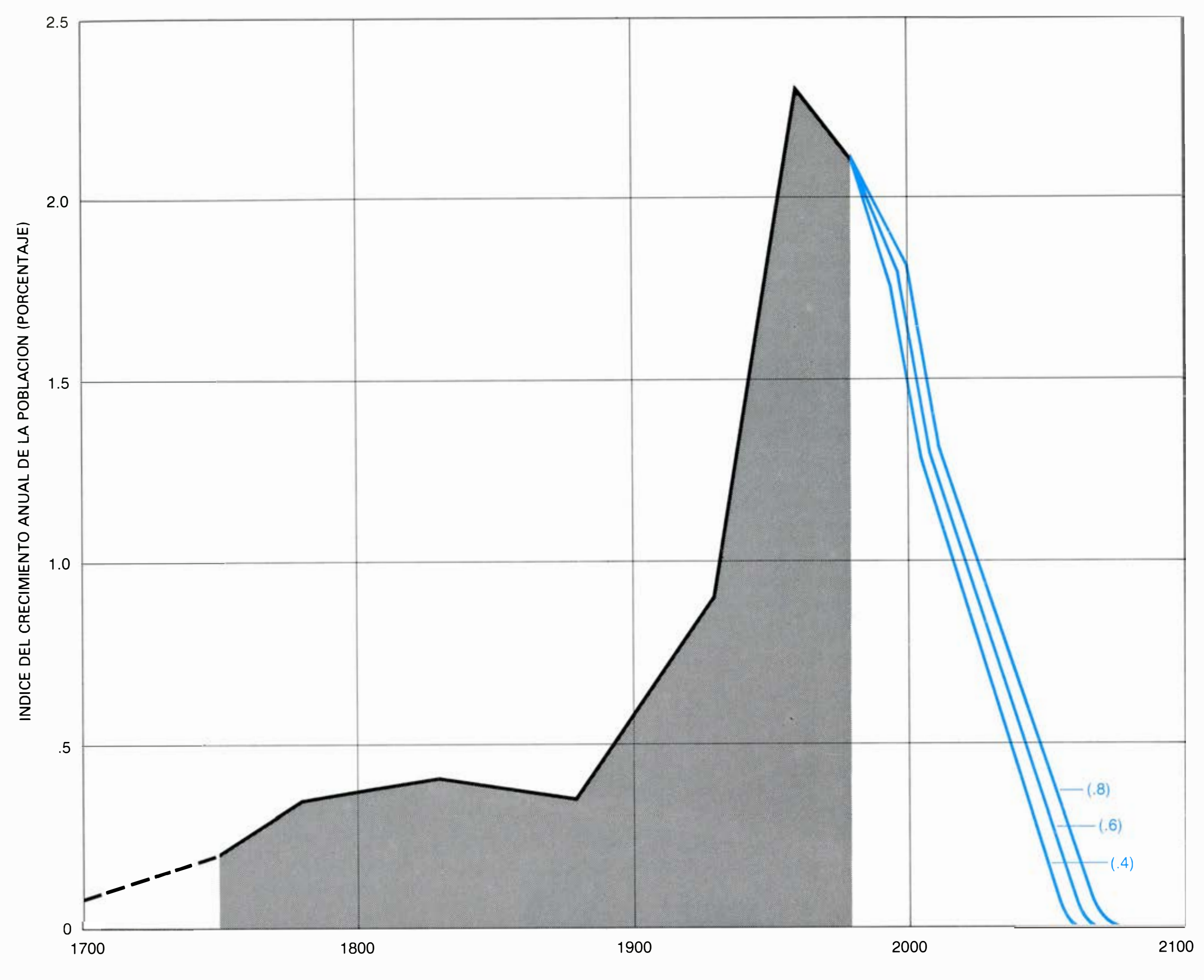
la población blanca de Sudáfrica, Taiwán y Estados Unidos. En general incluye un descenso un tanto más rápido de la mortalidad infantil que de la mortalidad adulta. Ha sido la familia más utilizada de las cuatro.

Las Naciones Unidas han formulado un modelo de edad ideal de descenso de la fecundidad para el Tercer Mundo y para cada una de sus regiones principales. La población inicial de los países en vías de desarrollo adoptada en nuestras simulaciones fue de 2890 millones en 1975. Las técnicas convencionales dan una proyección de la población del Tercer Mundo hasta el año 2000 a partir de estos supuestos y estimaciones. Una vez efectuada, se supone que la población en el año 2000 será de 4800 millones.

Con algunos supuestos adicionales pueden extenderse las proyecciones más allá del año 2000. Cabe presumir, por ejemplo, que la esperanza de vida continuará su ascenso hasta alcanzar un máximo y luego se mantendrá constante, y que la fecundidad bajará hasta un nivel suficiente para que la población se reponga a sí misma para luego estabilizarse también. Por este procedimiento se consigue un tamaño de población estable.

En tales proyecciones a largo plazo suele adoptarse como máximo para la esperanza de vida los 75 años, porque la experiencia de los países desarrollados muestra que constituye una cota accesible. Si la esperanza de vida continúa incrementándose a un ritmo de 0,4 años por año, alcanzará el máximo hacia el 2030. Si la fecundidad prosigue su descenso al ritmo supuesto para el período que va hasta el año 2000, llegará al nivel de reposición hacia el 2020 y se estabilizará al nivel mínimo de reposición de 2,1 hacia el 2030. La población detendrá su crecimiento hacia el 2100, con 8460 millones, casi tres veces la población de 1975. El gran aumento resulta del impulso anterior, así como de las tendencias de fecundidad y mortalidad proyectadas para el futuro.

El cálculo de población basado en esta extrapolación directa de experiencias recientes puede servir como línea de base con la cual comparar otras proyecciones. En concreto, podemos preguntarnos qué pasaría si la esperanza de vida se incrementara a mayor velocidad. El cálculo de la línea de base supone que la esperanza de vida se incrementará 0,4 años por año. Un ritmo tal implica que se interrumpirá la disminución de la pasada década, pero que no habrá un retorno al elevado nivel de los años cincuenta. Auque difícilmente descen-



INDICE DE CRECIMIENTO de la población de los países en vías de desarrollo, que alcanzó su cima en los años 60 y ha empezado ahora a descender. Los datos históricos se representan en la parte sombreada. (Sólo se dispone de datos fidedignos a partir de 1750.) El índice de crecimiento más alto se dio entre 1940 y 1975, cuando el descenso de mortalidad era el más rápido y la fecundidad todavía no había empezado a bajar. En su punto máximo, el índice de crecimiento del Tercer Mundo era aproximadamente de un 2,3 por ciento. Desde los años 60, las reducciones de la fecundidad marital han disminuido el índice hasta un 2,1 por ciento. Las reducciones futuras en el crecimiento se deberán principalmente a los cambios del índice de natalidad. Los aumentos de la esperanza de vida no tendrán un papel importante, como demuestra la concordancia de las curvas que representan los índices de crecimiento de población para los tres índices de mejora de la esperanza de vida. Bajo estos tres supuestos, el índice de crecimiento será nulo antes del 2100.

derá la mortalidad más deprisa que en esa década, las consideraciones antes expuestas inducen a pensar que, a través de un decidido esfuerzo, el ritmo de los años 50 podría alcanzarse otra vez. Un aumento anual de 0,6 años, un 50 por ciento más que el de la línea de base, se aproximaría a ese ritmo.

Si la esperanza de vida en el Tercer Mundo se incrementara regularmente 0,6 años por año, el máximo de esperanza de vida se alcanzaría hacia el 2015, 15 años antes que en el modelo de la línea de base. Admitiendo una misma población inicial, el mismo ritmo y modelo de descenso de la fecundidad y el mismo modelo de descenso de mortalidad que en la estimación de la línea de base, la población en el 2000 sería de 4900 millones, esto es, 100 millones, o el 2,1 por ciento, más que la estimación de la línea de base. Cuando la población alcanzara la estabilidad, hacia el 2100, sumaría 8940 millones de personas, unos 480 millones, un 5,7 por ciento, más que la población de la línea de base. La diferencia no es mayor debido al continuo descenso de la fecundidad y a causa también de que las muertes retrasadas corresponden progresivamente a personas mayores.

Si pudiera alcanzarse un incremento anual de la esperanza de vida de 0,8 años –el doble del ritmo actual– el incremento final de la población sería menos del 10 por ciento mayor que la población estable que supone la estimación de la línea de base. Ese ritmo nos llevaría a una población de 5000 millones en el 2000, un 4,2 por ciento mayor que la de base, y de 9160 millones en el 2100, un 8,3 por ciento mayor.

El efecto provocado por un incremento más rápido de la esperanza de vida variaría de una región a otra en los países en vías de desarrollo, en parte porque la esperanza de vida es mayor en unas áreas que en otras y, en parte también, porque la fecundidad es más baja, o se prevé que descienda más deprisa, en ciertas regiones. La mayor diferencia se daría en Africa, donde la fecundidad es más alta y, más baja, la esperanza de vida; estos valores no cambian allí tan rápidamente como en otras áreas. Un incremento anual en la esperanza de vida en Africa de un 50 por ciento más que el índice de la línea de base nos llevaría a una población un 9,2 por ciento mayor cuando el crecimiento se detenga. En Iberoamérica, la población en el 2100 sería un 1,8 por ciento mayor. En Asia oriental, la dife-

rencia se cifraría en un 1,6 por ciento y, en Asia meridional, de un 7,1.

Además del índice de aumento en la esperanza de vida hay otros aspectos de las tendencias de mortalidad futuras a tener en cuenta. Uno de ellos es la máxima esperanza de vida alcanzable. Otro, la distribución por edades del descenso de mortalidad. Sin embargo, probablemente ningún factor influya tanto en el crecimiento de la población como la tasa de aumento de la esperanza de vida. Si la esperanza de vida no rebasara en el Tercer Mundo nunca los 70 años, por ejemplo, la población estable sería menos de un 5 por ciento más reducida que el colectivo de la línea de base de 8460 millones. El propio descenso inmediato de la mortalidad infantil a los bajos niveles de Europa septentrional nos llevaría a una población estable que no llegaría a ser un 10 por ciento mayor que la de la línea de base. Puede imaginarse una combinación de los tres factores que afectara sustancialmente al tamaño de la futura población. Un conjunto tal de circunstancias vendría dado por un aumento lento (o incluso un descenso) de la esperanza de vida, un bajo nivel máximo de dicha esperanza y unos incrementos de ésta que incidieran casi exclusivamente en las reducciones de la mortalidad juvenil e infantil. En todo caso tal combinación cae, sin duda, fuera del orden de posibilidades consideradas probables por los expertos.

Comparado con el incremento causado por un rápido descenso del índice de mortalidad, el descenso resultante de una fuerte reducción de la fecundidad podría ser importante; supongamos que la fecundidad descendiera un 50 por ciento más rápido de lo que se establece en la estimación de la línea de base. En tal caso, de mantenerse la esperanza de vida con un aumento anual de 0,4 años, la población del año 2100 sería un 18,8 por ciento más pequeña. Ello representa más del triple que el 5,7 por ciento de diferencia causado por un aumento del 50 por ciento en el índice del incremento de la esperanza de vida de la línea de base.

Nuestro principal hallazgo hasta ahora (que la población del Tercer Mundo estabilizada sería un 6 por ciento mayor si la esperanza de vida aumentara un 50 por ciento más rápido) constituye sólo una aproximación tosca. Es probable que el aumento real de la población fuera menor, por la sencilla razón de que la mortalidad y la fecundidad no son independientes, contrariamente al supuesto simple empleado en nuestras proyecciones. Se dispone ya de bastante información, especialmente de los países en vías de desarrollo, en el sentido de que las fuertes reducciones de la mortalidad, infantil sobre todo, llevan a descensos más rápidos de la fecundidad.

La relación entre fecundidad y mortalidad infantil es particularmente fuerte en sociedades donde la mayoría de las mujeres crían a sus niños. La alimentación materna protege contra el embarazo inhibiendo la ovulación. La muerte de un niño elimina la protección y expone a la madre a la posibilidad de otro embarazo. Si una mayor proporción de niños sobreviven al período de alimentación materna, la fecundidad se reduce.

Además, muchos observadores afirman que, cuando la mortalidad infantil es alta, las parejas tienen más niños de los que desean, para asegurar que sobreviva el número elegido. Cuando la mortalidad desciende, el número de nacimientos necesario para alcanzar una determinada cifra de niños supervivientes también desciende. Crece el interés en limitar la familia y se reduce la fecundidad.

Aunque se haya probado la existencia de una tal relación, se discute aún su intensidad. Puesto que la población del Tercer Mundo crece con gran rapidez, el descenso de la fecundidad resultante de un bajo índice de mortalidad no compensa plenamente el crecimiento generado por la reducción de mortalidad. Como ha señalado Samuel H. Preston, de la Universidad de Pennsylvania, no habría problema de crecimiento demográfico en el Tercer Mundo si el efecto fuera totalmente compensador.

Sin embargo, en muchas situaciones cobra importancia un efecto parcial. En virtud del mismo, el incremento de un 5,7 por ciento en el tamaño de población estable, que resulta de elevar el índice de aumento de la esperanza de vi-

EFECTO DE LA FECUNDIDAD sobre el tamaño de la población del Tercer Mundo en el punto de estabilización, hacia el año 2100; probablemente sea mucho mayor que el efecto de la mortalidad. Las barras muestran la diferencia entre la población de la línea de base y la población que resultaría de un descenso más rápido de la mortalidad o de la fecundidad. Si la esperanza de vida creciera un 50 por ciento más rápidamente de lo que ahora crece, la población estable (*color oscuro*) sería aproximadamente un 6 por ciento mayor que la población de la línea que sirve de base. Si la fecundidad bajara un 50 por ciento más rápidamente de lo que supone la estimación de la línea de base, la población estable (*gris claro*) sería un 19 por ciento menor que la de la línea de base. Si doblamos los dos índices, los resultados son análogos (*color claro* y *gris oscuro*). El efecto de un tal descenso de la fecundidad sobre el tamaño de la población futura parece representar el triple escaso del efecto de un descenso más rápido de la mortalidad.

da de unos 0,4 a un 0,6, supone, casi con seguridad, el punto máximo de cualquier grupo razonable de estimaciones. A la luz del trabajo efectuado hasta ahora, sobre la relación entre mortalidad y fecundidad, una diferencia de un 3 a un 4 por ciento probablemente se ajuste mejor.

Ni esta estimación de la población adicional ni cualquier otra deben considerarse inevitables. El volumen del incremento de población provocado por un descenso más rápido de la mortalidad no puede determinarse. Variará según las medidas políticas que hayan contribuido a producirlo. Según sean esas medidas, la mortalidad afectará de una forma u otra a los distintos grupos de edad. Por las razones tratadas más arriba, los programas de reducción de la mortalidad juvenil tendrían un impacto algo mayor que los que afectan principalmente a los ancianos. Ni siquiera este efecto reviste la importancia que da a entender. Como hemos señalado antes, la población estable definitiva del Tercer Mundo seguiría superando el colectivo de la línea de base en una cifra inferior al 10 por ciento, aunque se eliminaran inmediatamente la mayoría de las muertes infantiles. Considerando la influencia, sobre la fecundidad, de un rápido descenso de la mortalidad, la diferencia giraría en torno al 5 por ciento.

Mucho más significativa es la influencia sobre la fecundidad de los programas concretos por los cuales se reduce la mortalidad. Se presume que algunas medidas ejercerán sólo un efecto relativamente pequeño e indirecto sobre la fecundidad. Si la mortalidad infantil se reduce gracias a la inmunización o a la administración de antibióticos, por ejemplo, puede esperarse que la fecundidad baje, a medida que las parejas comprendan que necesitan menos nacimientos para alcanzar un número dado de hijos. Con más hijos supervivientes habría más crianza materna, lo que haría descender también la fecundidad. De cualquier manera, el descenso de la fecundidad causado por tales cambios sería, probablemente, bastante lento.

Otras medidas pueden ejercer un efecto más fuerte y directo. La educación de la mujer, por ejemplo. Se ha demostrado que el aumento de la alfabetización y la promoción de la mujer en la sociedad rural produce un descenso de la mortalidad juvenil e infantil. Ese descenso, a su vez, motivaría una reducción de la fecundidad análoga a la de un programa de inmunización. Al mismo tiempo, los cambios resultantes de una mejora en la educación podrían tener, en muchos aspectos, una influencia adicional sobre la fecundidad. Aunque hay excepciones, en muchas sociedades la promoción de la mujer despierta aspiraciones, para ellas y para sus hijos. Las aspiraciones, a menudo, se realizan mejor retrasando el matrimonio, distanciando los nacimientos y reduciendo el número de hijos.

Los servicios de planificación familiar son otro ejemplo. La planificación familiar sirve para reducir la mortalidad infantil, capacitando a las parejas para prolongar los intervalos entre hijos, lo que produce una crianza más sana, y menguando el número adicional de nacimientos en el caso de las mujeres de edad con muchos hijos. (El índice de mortalidad es especialmente alto entre los niños nacidos en esas circunstancias.) En ese sentido, la planificación familiar influye directamente sobre la fecundidad.

Poner el énfasis en medidas de esta índole, a la hora de concentrar los esfuerzos por atajar la mortalidad, conllevaría algún tipo de coste. No puede determinarse con certeza en este momento hasta qué punto se vería afectado el crecimiento por ello. Por lo que se sabe, no parece ilusorio creer que una planificación esmerada produzca un paquete de medidas pensadas para reducir la mortalidad, lo cual compensaría los descensos de fecundidad, o incluso medidas de más vasto alcance.

Si caemos en la cuenta de que una diferencia de sólo un 5 por ciento en la población de los países en vías de desarrollo en el 2100 representará entre 400 y 500 millones de personas, constituiría un gran adelanto el que los programas para incrementar la esperanza de vida se diseñaran de suerte tal que minimizaran sus consecuencias sobre el crecimiento de la población. Aun cuando así no fuera, lo que llevamos dicho sugiere que, en el futuro, el ritmo de descenso de la mortalidad es probable que no sea el factor principal que influya sobre el índice de crecimiento de la población. Tendencias pasadas han configurado un proceso que nos lleva a una población del Tercer Mundo aproximadamente el triple de la actual, antes de alcanzar la estabilización. Reducciones más rápidas de la fecundidad podrían cambiar un tanto esta perspectiva, pero reducciones más rápidas en la mortalidad no tendrían más que un impacto limitado. Las reservas acerca de los esfuerzos para reducir más rápidamente la mortalidad, argumentando un supuesto crecimiento de la población, están, por tanto, fuera de lugar.

Quarkonio

Un "átomo" formado por un quark y un antiquark pesados integra el mejor sistema disponible para examinar las fuerzas que ligan entre sí a los constituyentes elementales de las partículas subnucleares

Elliott D. Bloom y Gary J. Feldman

El contexto más adecuado para investigar las fuerzas de la naturaleza es un sistema de dos objetos ligados entre sí por atracción mutua. La Tierra y la Luna, por ejemplo, constituyen el sistema más directamente accesible para observar la fuerza gravitatoria. El átomo de hidrógeno, que consta de un electrón y un protón, ha venido siendo el banco de pruebas básico donde someter a contrastación las teorías de la fuerza electromagnética. El deuterón, formado por un protón y un neutrón, constituye el sistema modelo para estudiar las fuerzas del núcleo atómico. Sabemos ya de un sistema ligado donde estudiar la fuerza que actúa entre los quarks, es decir, los constituyentes de los protones, neutrones y otras partículas similares. El sistema se denomina quarkonio (o quarkonium): consta de un quark pesado ligado a otro quark igualmente pesado. La fuerza que interviene en el quarkonio es la más fuerte de las conocidas. Bautizada como "fuerza de color", se cree que constituye el fundamento de todas las fuerzas nucleares.

El positronio, un átomo artificial, es, en algunos aspectos, el sistema más sencillo de los formados por dos cuerpos. Consta de un electrón ligado a un positrón (la antipartícula del electrón). Al igual que el átomo de hidrógeno, el positronio se mantiene ligado por la atracción de cargas eléctricas opuestas, aunque tiene una estructura más simétrica. En efecto, el protón del átomo de hidrógeno es mayor y más pesado que el electrón; no ocurre esto con las dos partículas que componen el positronio, que son idénticas en tamaño y masa. Se han observado varios estados de positronio; en cada uno de ellos, el electrón y el positrón poseen un modo de movimiento único y, por ende, cada estado ostenta una energía definida. A partir del espectro de energías de los estados se deducen determinadas características de la fuerza electromagnética.

La estructura del quarkonio es muy similar a la del positronio. En el quarkonio, el quark y el antiquark, ligados, son también idénticos en tamaño y masa; asimismo, los modos de movimiento permitidos son iguales a los del positronio, por lo que puede esperarse un espectro similar de estados de energía. A lo largo de los últimos años hemos asistido a un empeño decidido por detectar los estados del quarkonio; fruto de ello, se han encontrado una docena, en números redondos. Desde un punto de vista cualitativo, tales estados se corresponden con los del positronio. En otro orden, el sistema quarkonio es la cienmilésima parte en tamaño del sistema positronio, pero la masa o energía total del primero es de 3000 a 10.000 veces mayor que la del segundo. El examen del espectro del quarkonio nos revela que la fuerza de color no sólo es más fuerte, sino también más compleja que la electromagnética.

Los quarks entraron en el mundo de la física, en 1964, de la mano de Murray Gell-Mann y George Zweig, del Instituto de Tecnología de California. El modelo quark pretendía explicar la diversidad de los hadrones, partículas que hasta entonces se habían considerado últimas o elementales. Los quarks se combinan de dos maneras distintas para formar un hadrón. Tres quarks ligados constituyen un barión, miembro de la subclase de los hadrones que engloba al protón y al neutrón. El sistema integrado por un quark y un antiquark ligados origina un mesón. El más ligero y común de los mesones es el pion; los sistemas quarkonio de los que vamos a ocuparnos son también mesones, aunque mucho más pesados.

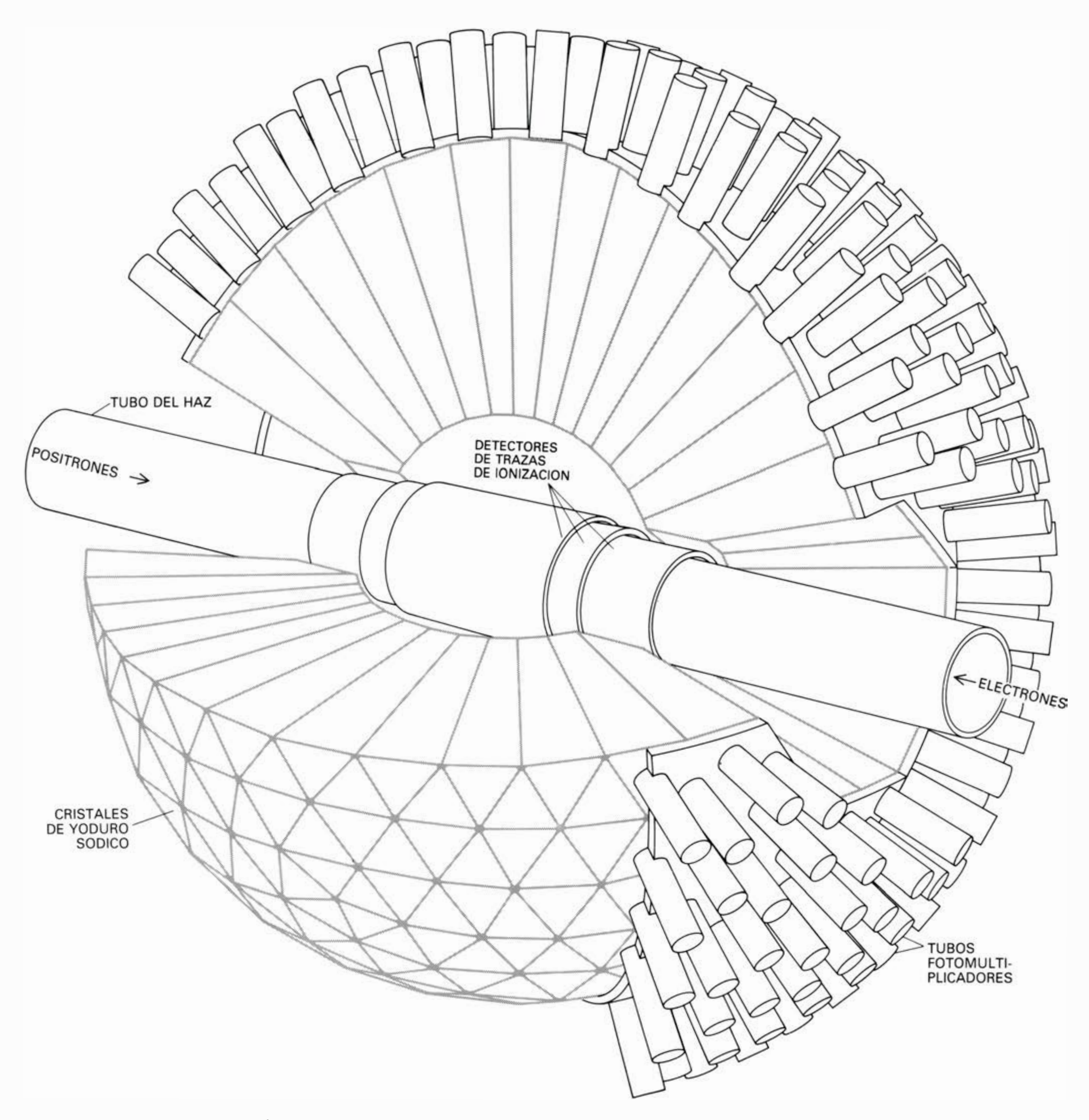
Se creyó al principio que sólo habría tres tipos de quarks: "up" (hacia arriba), "down" (hacia abajo) y "strange" (extraño), conocidos por sus iniciales *u*, *d* y *s*. Todos los hadrones de los que se tenía noticia podían interpretarse como combinaciones de esos tres quarks (y sus antiquarks correspondientes $\bar{u}$, $\bar{d}$ y $\bar{s}$). Veíase así que el sistema *uud* tenía todas las propiedades del protón y que el sistema $u\bar{d}$ podía identificarse como el pion cargado positivamente. Se diría que esta rica variedad de hadrones ofrecía amplias posibilidades para explorar la fuerza entre quarks. La verdad es que los experimentos realizados con protones, neutrones, piones y otros hadrones "ordinarios" sólo proporcionan información indirecta acerca de la fuerza entre quarks. La razón estriba en que los quarks *u*, *d* y *s* son bastante ligeros; su masa, expresada en unidades de energía, se parece a la energía de ligadura que mantiene unidos a los quarks en el hadrón. En virtud de ello, en un hadrón ordinario los quarks se mueven a velocidades próximas a la de la luz; para calcular sus propiedades se recurre a métodos complejos de la teoría especial de la relatividad. Cálculos que encierran una dificultad excesiva para resultar prácticos.

Lo que se necesitaba era un sistema ligado de quarks pesados, en los que la energía de ligadura fuera pequeña comparada con la masa del quark. Los quarks se moverían entonces a velocidades mucho menores que la de la luz y podrían dejarse de lado, entonces, las complicaciones de la teoría de la relatividad. En 1974 se dio con un sistema de quarks no relativista de este tipo, al descubrirse un extraordinario mesón dotado de una masa (en unidades de energía) de 3095 millones de electronvolt (MeV). Hallaron la partícula, casi simultáneamente, dos grupos experimentales adscritos, respectivamente, al Laboratorio Nacional de Brookhaven y al Stanford Linear Accelerator Center (SLAC). El grupo de Brookhaven denominó a la nueva partícula *J*; ψ (psi griega) la llamaron los de Stanford, y con ese nombre la usaremos nosotros aquí.

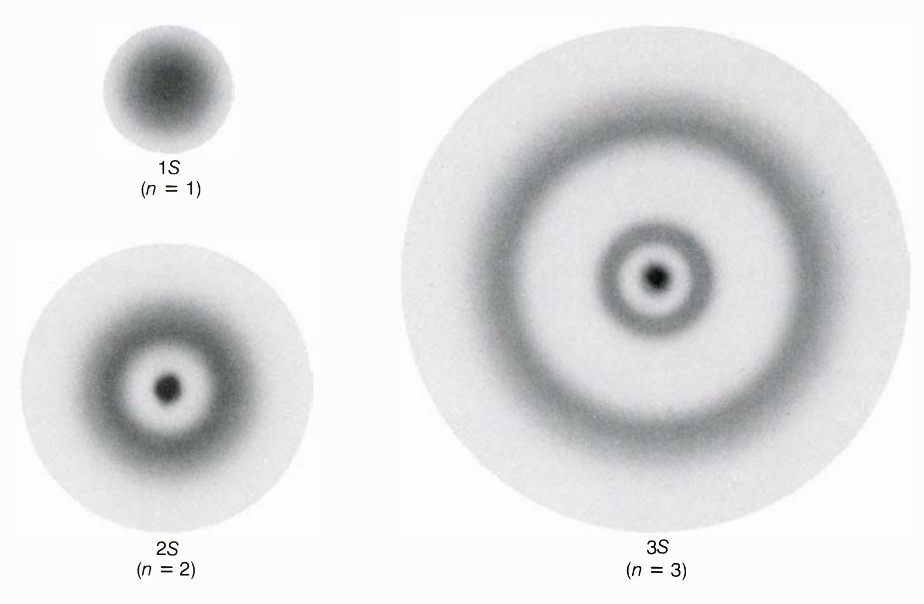
¿Qué es el mesón ψ? No parecía probable que se tratara de una combina-

ción de los quarks u, d y s; en parte, porque todas las combinaciones con las propiedades adecuadas correspondían a los hadrones ya conocidos. Unos 10 años antes, sin embargo, James D. Bjorken, a la sazón en el SLAC, y Sheldon Lee Glashow, de la Universidad de Harvard, habían especulado sobre la posible existencia de un cuarto tipo o "sabor" de quark, al que caprichosamente denominaron "charm" o "encanto" (c). En 1970, Glashow y sus colegas John Iliopoulos y Luciano Maiani argumentaron, desde supuestos teóricos, que el quark "encantado" debía existir y ser sustancialmente más pesado que los quarks restantes. Poco antes del descubrimiento de la ψ, Thomas W. Appelquist y H. David Politzer, que por entonces se hallaban también en Harvard, habían señalado que un quark y un antiquark encantados podrían formar un estado ligado no relativista. Y bautizaron a ese estado "charmonio", a semejanza del positronio. La partícula ψ se reinterpretó de inmediato como una forma de charmonio: un mesón integrado por los quarks $c\bar{c}$.

El descubrimiento del charmonio estimuló la búsqueda de quarks más pe-



DETECTOR DE LA BOLA DE CRISTAL, uno de los instrumentos empleados para explorar la estructura del quarkonio, el estado de materia formado por un quark y un antiquark del mismo tipo. El detector está instalado en un anillo de almacenamiento de partículas; haces de alta energía de electrones y positrones (o antielectrones) colisionan en el centro del detector y generan nuevas partículas, entre ellas el quarkonio. El quarkonio se desintegra entonces y produce más partículas, cuyo paso queda registrado. La Bola de Cristal posee una extrema sensibilidad a los fotones o cuantos de la radiación electromagnética. Los elementos principales del detector son 732 cristales de yoduro sódico que rodean la zona de colisión con una disposición basada en la simetría de un icosaedro. (No se muestran los cristales adicionales de los extremos del detector.) El cristal que absorbe un fotón emite luz que puede registrarse por un tubo fotomultiplicador. La dirección del fotón y la cantidad de energía que deposita en el cristal proporcionan información sobre el estado del quarkonio que lo ha emitido. Los detectores de trazas de ionización ayudan a caracterizar las partículas cargadas. La Bola de Cristal ha funcionado durante tres años en el SPEAR, un anillo de almacenamiento del SLAC. Pero se la ha trasladado al DORIS, un anillo de mayor energía del DESY alemán.



ESTADOS DE UN SISTEMA LIGADO formado por una partícula y una antipartícula. Se describen mediante funciones de onda, que dan la probabilidad de encontrar a la partícula en un punto cualquiera del espacio. La figura muestra la distribución de densidad de probabilidad para tres estados de positronio, el sistema ligado constituido por un electrón y un positrón. Las regiones más oscuras de las tres figuras corresponden a las de mayor probabilidad. Cada estado se identifica por su número cuántico principal, *n*, y por una letra que denota el momento angular orbital del sistema. Las tres configuraciones corresponden a estados *S* cuyo momento angular orbital es cero. El quarkonio forma estados análogos.

sados todavía. Había razones para creer que se presentarían a pares. Los dos inmediatos en el elenco de los ya conocidos, se postulaba, serían "bottom" (o "fondo", *b*) y "top" (o "cima", *t*). En 1976, Estia Eichten y Kurt Gottfried, de la Universidad de Cornell, sugirieron que el "bottomonio" (el mesón con el contenido en quarks $b\bar{b}$) y el "toponio" ($t\bar{t}$) deberían formar sistemas no relativistas igual que el charmonio, si bien portarían un espectro de estados ligados mucho más rico. El primer estado del bottomonio se descubrió en 1977, en el Fermi National Accelerator Laboratory (Fermilab), instalado en los aledaños de Chicago. Se le puso el nombre de Y (hypsilon griega); su masa es de 9460 MeV. Falta por detectar el sistema toponio, que, si existe de verdad, tendrá más de 36.000 MeV de masa.

A lo largo de los últimos ocho años se ha estudiado hasta sus pormenores el espectro del charmonio. Se ha empezado la exploración del bottomonio. Además, se han identificado varias partículas formadas por un quark "encantado" en combinación con un quark de otro tipo; de tales partículas se dice que exhiben "encanto" neto. Acaba de informarse de ciertas pistas que confimarían el "fondo" neto. Aquí nos ceñiremos principalmente a los sistemas en los que un quark pesado está ligado a un antiquark del mismo tipo, como el $c\bar{c}$, $b\bar{b}$ y el conjeturado $t\bar{t}$. A esos sistemas les damos el nombre genérico de quarkonio.

Para entender qué fuerzas actúan en el quarkonio me parece pedagógico empezar con un repaso del átomo de hidrógeno y el positronio. La ley de Coulomb describe la fuerza más importante que liga a un electrón con un protón o con un positrón. Dice esta ley que esa fuerza es directamente proporcional al producto de las cargas eléctricas de las partículas e inversamente proporcional al cuadrado de la distancia que las separa. La constante de proporcionalidad, α, es una medida de la intensidad inherente de la fuerza electromagnética; su valor numérico viene a ser 1/137.

De acuerdo con la física clásica, las partículas del positronio o del átomo de hidrógeno deberían acercarse hasta colisionar. Pero la experiencia no corrobora la predicción. Lo que se convirtió en un espinoso problema en los primeros intentos por formular la teoría del átomo sólo se resolvió con el desarrollo de la mecánica cuántica en la década de 1920. El modelo del átomo de hidrógeno propuesto por Niels Bohr en 1913 constituyó un paso importante en este sentido. Bohr postuló que el electrón sólo puede ocupar ciertas órbitas discretas en el átomo de hidrógeno. En cada órbita permitida, el electrón tiene una energía definida; y únicamente cambiará su energía si salta bruscamente a otra órbita permitida. Al cambiar de órbita, emite o absorbe un fotón, esto es, un cuanto de radiación electromagnética, cuya energía coincide con la diferencia entre las energías de las dos órbitas.

Bohr demostró que la energía de ligadura de un electrón en una órbita dada era igual a E_0/n^2, donde E_0 es la energía de ligadura correspondiente a la órbita más pequeña y *n* es un número entero positivo llamado número cuántico principal. La energía de ligadura es la que debe suministrarse para separar completamente al electrón del protón. Para el átomo de hidrógeno, E_0 se cifra en 13,6 electronvolt. Los valores de *n* se asignan correlativamente a las órbitas, empezando por la más pequeña. Para ésta, *n* es 1 y, 13,6 electronvolt, la energía de ligadura; para la órbita siguiente, *n* es 2 y la energía de ligadura es de $13{,}6/2^2$ (es decir 3,4) electronvolt; y así sucesivamente.

El modelo de Bohr puede aplicarse al positronio. Mas, debido a la diferencia en masa entre el protón y el positrón, el radio de una órbita del positronio es el doble que el de la órbita de un átomo de hidrógeno con el mismo número cuántico principal. La energía de ligadura del positronio es, por tanto, la mitad: 6,8 electronvolt para la órbita más pequeña. Aunque la secuencia de estados es la misma, todas las energías están reducidas en un factor dos.

La descripción completa de un estado ligado exige especificar el momento angular, amén de la energía. En un sistema de dos cuerpos intervienen tres elementos en la configuración del momento angular total: la rotación de cada uno de los cuerpos alrededor de su propio eje y el movimiento orbital de los cuerpos alrededor de su centro de masa. Cada una de estas cantidades puede representarse por un vector orientado según el eje de rotación y cuya longitud es igual a la magnitud del momento angular. El momento angular de un sistema de cuerpos podría describirse mediante la magnitud y la dirección de los tres vectores. Hay, sin embargo, otros métodos de representación que, sin menoscabar la información, reflejan mejor las simetrías del sistema. Para describir los estados del positronio y del quarkonio resulta más ilustrativo especificar el momento angular orbital, la suma de los dos momentos angulares intrínsecos o de espin y el momento angular total.

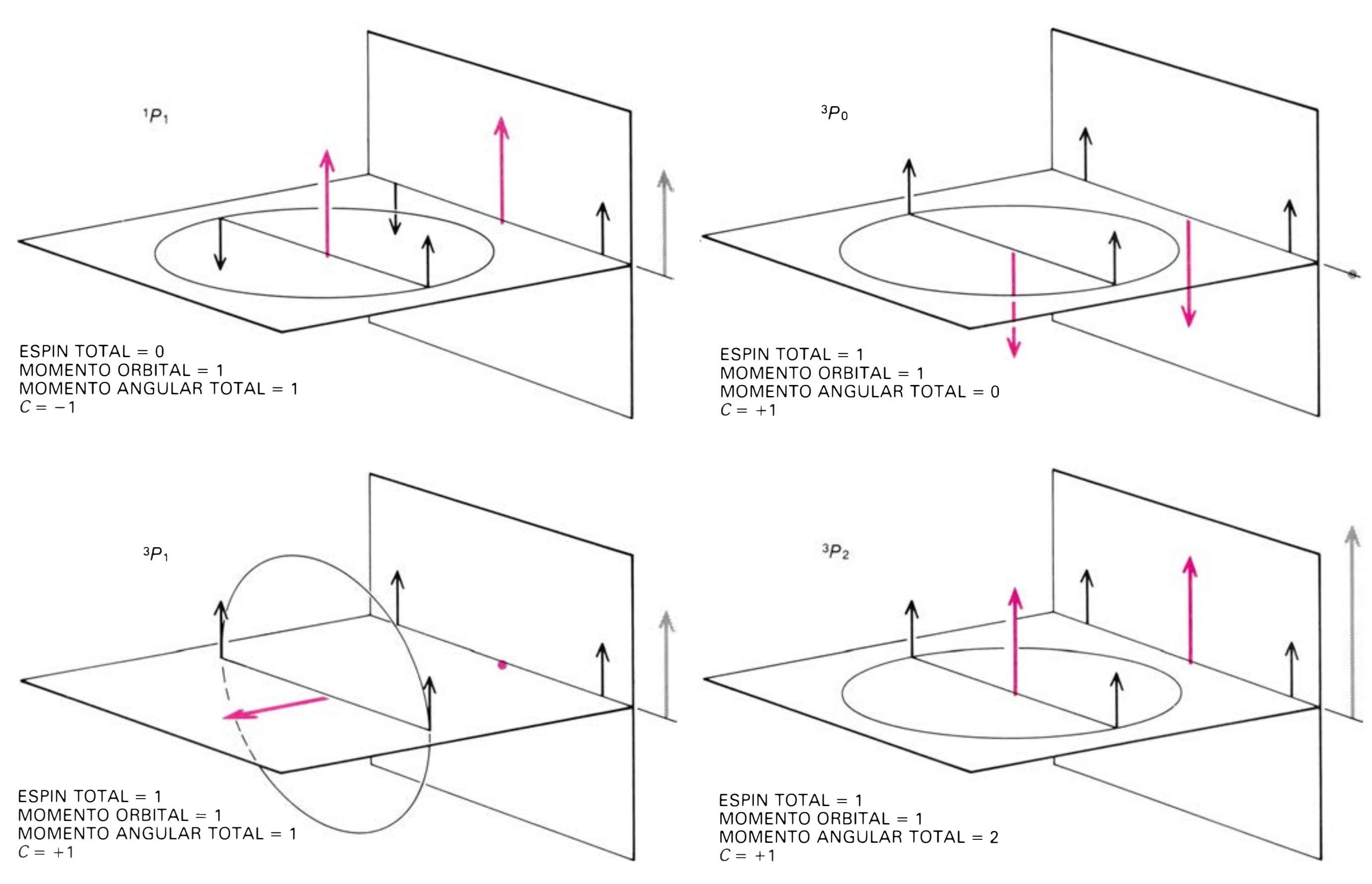
En un sistema a gran escala donde pueda prescindirse de las leyes de la mecánica cuántica, los vectores de momento angular tendrán cualquier magnitud y orientación. Más sencilla resulta, en algunos aspectos, la descripción del momento angular en un sistema cuántico. La mecánica cuántica impide

que un observador llegue a un conocimiento exacto de un vector de momento angular; sólo puede medirse la magnitud absoluta del vector y su proyección sobre un eje en el espacio (que puede escogerse libremente). Además, sólo se admiten ciertos valores discretos para estas cantidades: un momento angular ha de ser entero o semientero cuando se expresa en unidades de la mínima cantidad posible de momento angular (que es igual a la constante de Planck dividida por 2π). Otra restricción: el momento angular orbital ha de ser siempre menor que n, el número cuántico principal.

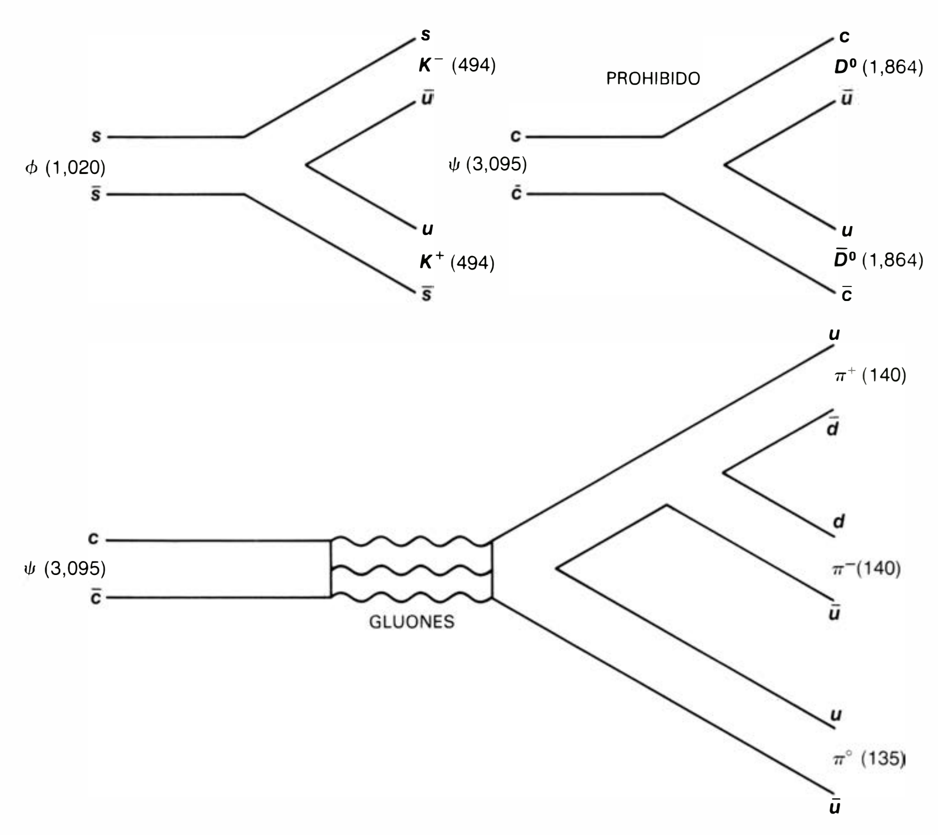
El momento angular orbital sólo puede tomar valores enteros 0, 1, 2, etcétera. En una notación compacta, introducida para describir los espectros atómicos, algunos años antes de que Bohr desarrollara su modelo, los sucesivos valores del momento angular orbital se designan por las letras *S, P, D,* y *F.* La órbita de Bohr y el momento angular de un estado pueden especificarse, por tanto, añadiendo la letra apropiada al número cuántico principal. Así, el estado de mínima energía del positronio se designa por $1S$; corresponde al estado en que $n = 1$ y carece de momento angular orbital. El símbolo $2P$ designa un estado con $n = 2$ y una unidad de momento angular orbital.

Tanto en el positronio como en el quarkonio, las distintas partículas constituyentes tienen su propio momento angular intrínseco (o espín) de 1/2 unidad. Los espines pueden combinarse de dos maneras. Si apuntan en direcciones opuestas, se cancelan mutuamente y el sistema tendrá un espín total igual a cero. Si son paralelos, se suman y el espín total valdrá 1. El momento angular total de un sistema depende de la orientación relativa del momento angular orbital y el espín total; más concretamente, depende de la proyección del vector de espín total sobre la dirección del vector orbital. Cuando el espín total valga cero, la única proyección posible será obviamente nula. A tal estado se le llama singlete, ya que para un determinado valor del momento angular orbital existe un único valor posible del momento angular total. Cuando el espín total es 1, tiene tres posibles proyecciones sobre el eje orbital: $+1$, 0 y -1. Estos valores corresponden a las tres posibles orientaciones: paralela, perpendicular y antiparalela. Un estado de este tipo recibe el nombre de triplete, ya que para un determinado valor del momento angular orbital existen tres estados distinguibles que difieren en el momento angular total.

El espín y el momento angular totales se incorporan a la notación de los estados cuánticos a través de un sistema de superíndices y subíndices. Para estados singletes se escribe el superíndice 1 delante de la letra que designa al estado; para estados tripletes el superíndice es 3. El momento angular total viene dado por un subíndice añadido al símbolo. Así, 2^1P_1 designa el estado del



PODEMOS ANALIZAR EL MOMENTO ANGULAR de un sistema ligado por medio de vectores proyectados sobre un plano. En el momento angular intervienen dos factores, a saber: el espín de las partículas y su movimiento orbital. Los electrones y los quarks tienen un espín intrínseco de 1/2 (*flechas negras*); en los estados mostrados aquí, el momento orbital es igual a 1 (*flechas de color*). Si las proyecciones de los vectores de espín apuntan en direcciones opuestas, se anula el espín total del sistema y el momento angular total (*flechas grises*) se hace igual al momento orbital. El estado resultante se denomina singlete, debido a que las proyecciones del espín y del momento orbital sólo pueden tener una suma posible. El estado se designa por $^1P_1$, donde P significa una unidad de momento angular orbital, el superíndice 1 identifica al estado como singlete y el subíndice 1 designa el momento angular total. Si los vectores de espín son paralelos, el espín total vale 1; éste puede tener tres posibles orientaciones con respecto al momento orbital. Existe, por tanto, un triplete de estados designados $^3P_0$, $^3P_1$ y $^3P_2$. El superíndice 3 denota triplete y los subíndices 0, 1 y 2 dan el momento angular total cuando las proyecciones de los vectores espín total y de momento orbital son respectivamente antiparalelas, perpendiculares y paralelas. El número de conjugación de carga (C), otra propiedad del estado ligado, se determina sumando, en vez de las proyecciones, las magnitudes absolutas del espín total y el momento angular orbital. Cuando la suma es un número impar (como en el estado $^1P_1$), C tiene el valor -1; cuando la suma es par (como en los estados ^3P), C es entonces igual a $+1$.



LOS SISTEMAS QUARK-ANTIQUARK se clasifican en ligados o cuasi-ligados, según el modo de desintegrarse. La partícula denominada mesón φ (phi) es cuasi-ligada. Conforma el estado 1^3P_1 de un quark "extraño" *s* y un antiquark "extraño" $\bar{s}$; su masa es de 1020 millones de electronvolt (MeV). Cuando el φ se desintegra, el *s* y el $\bar{s}$ se separan y se constituye un nuevo par quark-antiquark; los productos son los mesones denominados K^+ y K^-, con una masa conjunta de 988 MeV. Los 32 megaelectronvolt restantes se convierten en energía cinética. Para el mesón ψ (psi), el proceso de desintegración análogo está prohibido por la conservación de la energía. El ψ corresponde al estado 1^3S_1 de un quark "encantado" *c* y un antiquark "encantado" $\bar{c}$ y su masa es de 3095 MeV. Las partículas más ligeras que contienen quarks encantados son los mesones D^0 y $\bar{D}^0$, cuya masa conjunta es de 3728 megaelectronvolt. La creación de este par a partir del mesón ψ conllevaría un déficit de 633 MeV. El ψ se desintegra por aniquilación en tres gluones, dando lugar a pares quark-antiquark más ligeros y, en última instancia, a hadrones ordinarios. El mesón ψ constituye un estado ligado de materia denominado charmonio, análogo al positronio.

positronio con n igual a 2, una unidad de momento angular orbital y espín total igual a cero. Al no haber espín total, el estado es un singlete y el momento angular total es necesariamente igual al momento angular orbital. El triplete de estados 2^3P_0, 2^3P_1 y 2^3P_2 tiene el número cuántico principal y el momento angular orbital común y el espín total es 1. Los tres estados se distinguen por la orientación del espín total con respecto al eje orbital, dando lugar a los tres valores del momento angular total.

La fuerza de Coulomb entre un electrón y un positrón depende sólo de las cargas de estas partículas y de la distancia entre ambas; es, pues, independiente del momento angular. Si la fuerza de Coulomb fuese la única que actúa en el positronio, todos los estados de momento angular con el mismo valor de n tendrían la misma energía. Pero existen otras fuerzas. Tanto el movimiento orbital como el de espín generan campos magnéticos que, a su vez, causan atracciones y repulsiones que pueden alterar la energía del estado. Por ejemplo, en el estado 1^1S_0 los dos espines son antiparalelos y, atractiva, la interacción magnética resultante. En el estado 1^3S_1 los espines son paralelos y la interacción magnética, repulsiva. Como consecuencia de ello la energía del primer estado es ligeramente inferior a la del segundo.

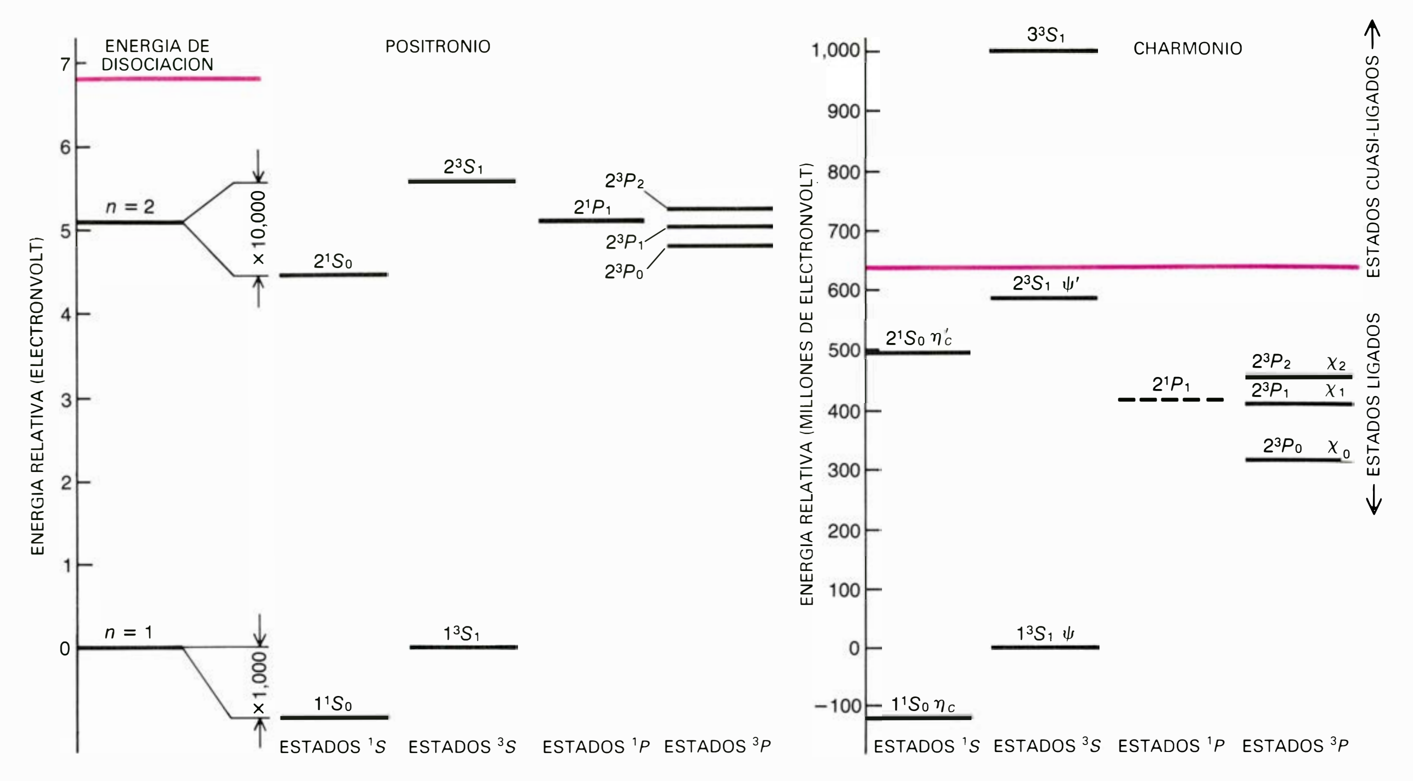
En el positronio, el incremento de energía asociado al momento angular de las partículas constituyentes es, como mucho, 5000 veces menor que la energía de ligadura. En el quarkonio pueden observarse efectos análogos: el movimiento orbital y el espín generan campos magnéticos de color. Las diferencias de energía son, sin embargo, mucho mayores. La energía de ligadura del quarkonio supera, 100 millones de veces, la del positronio; la separación típica entre estados con distinto momento angular es de un quinto de la energía de ligadura.

La analogía entre el positronio y el quarkonio puede extenderse a un nuevo fenómeno: la desintegración de los sistemas ligados. Cuando un electrón y un positrón se juntan, se aniquilan mutuamente; la energía equivalente a su masa aparece entonces en forma de radiación electromagnética. Un quark y un antiquark pertenecientes a la misma clase pueden aniquilarse entre sí de manera similar, aunque ahora la energía toma inicialmente la forma de radiación asociada a la fuerza de color. Mas para la descripción de estos procesos no nos sirve ya el modelo atómico de Bohr.

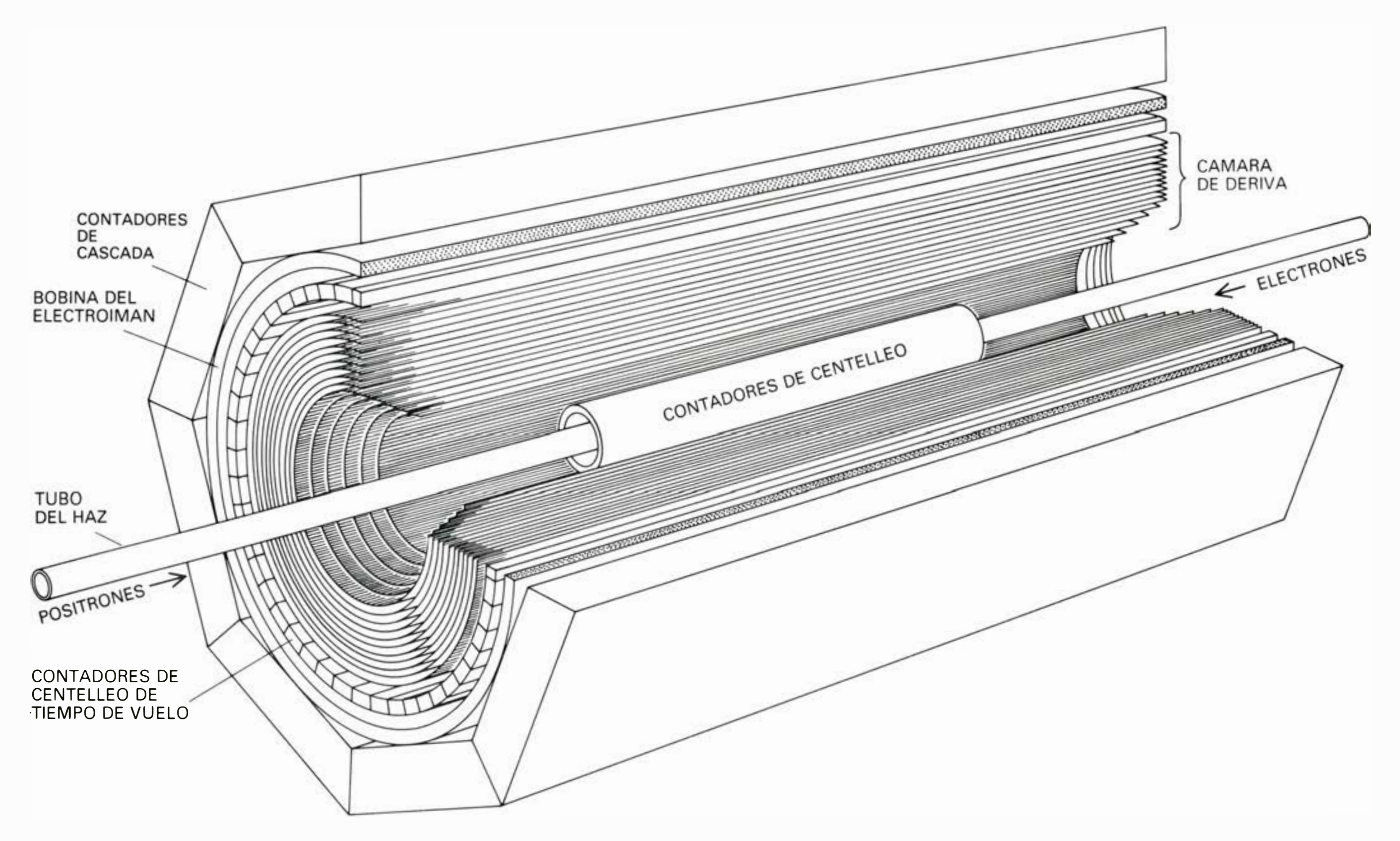
Se ha sustituido el modelo en cuestión por una versión más refinada de la mecánica cuántica. En ese nuevo marco, el electrón y el positrón carecen de una órbita definida, ni se les puede asignar una posición precisa en un instante determinado. Lo único que puede calcularse es la probabilidad de encontrar una partícula en un punto dado del espacio. Esta probabilidad viene definida por una función matemática llamada función de onda, que es distinta para cada estado. Para los estados *S* del positronio, la probabilidad de desintegración por aniquilaciám mutua es proporcional a la probabilidad de encontrar las dos partículas en el mismo punto.

La desintegración está gobernada por varias leyes de conservación. Así, la carga eléctrica sumada de todas las partículas después de la desintegración ha de ser igual a la que había antes de la desintegración. Debe mantenerse, asimismo, la energía y el momento lineal. La energía de un sistema positronio es aproximadamente de 1 MeV; el momento lineal posee un valor nulo, ya que podemos considerar que las partículas se mueven en direcciones opuestas cuando colisionan. El fotón puede acarrear cualquier cantidad de energía, pero al ser su masa igual a cero, su momento lineal tiene que coincidir, invariablemente, con su energía. Si el positronio fuera a desintegrarse en un único fotón, la energía y el momento lineal no podrían conservarse a la vez. Por esta razón, la desintegración siempre da lugar a dos fotones, por lo menos. Si se emiten los dos fotones en direcciones opuestas, cada uno acarrea la mitad de la energía del sistema y sus momentos lineales se cancelan.

Una propiedad de la función de onda llamada número de "conjugación de carga" nos da una restricción adicional sobre el número de fotones emitidos. La conjugación de carga es una operación imaginaria por la que se convierten todas las partículas en sus antipartículas. Cuando esta operación se aplica al positronio, el electrón se convierte en positrón y el positrón en electrón.



ESPECTRO DE ESTADOS DE ENERGIA en el positronio y en el charmonio: similar en ambos. La escala de las diferencias de energía es mayor en el charmonio en un factor de aproximadamente 100 millones. La energía de un estado viene determinada por el número cuántico principal *n* y por la orientación de los espines de las partículas y del momento angular orbital. En el positronio, las distintas combinaciones de momento angular sólo causan minúsculos corrimientos en la energía (mostrados con la escala vertical expandida), pero en el charmonio los corrimientos son mucho mayores. Todas las energías están dadas con relación al estado 1^3S_1. El positronio se disocia a 6,8 electronvolt. A 633 megaelectronvolt por encima de la energía del mesón ψ el charmonio se convierte en cuasi-ligado, ya que puede desintegrarse en los mesones D^0 y $\bar{D}_0$. El estado 2^1P_1 no se ha detectado experimentalmente.



DETECTOR MARK II del anillo de almacenamiento SPEAR. Se muestra muy eficaz a la hora de medir el momento de partículas cargadas eléctricamente. La mayor parte de su volumen corresponde a una cámara llena de gas, donde se disponen, en cilindros concéntricos, unos 12.800 alambres cargados. Si un partícula cargada atraviesa la cámara, ioniza átomos del gas; los electrones liberados son atraídos hacia un alambre dotado de carga positiva, originando una corriente que puede medirse electrónicamente. La cámara está inmersa en un campo magnético que obliga a la partícula cargada a seguir una trayectoria curva; a partir del radio de curvatura puede medirse el momento de la partícula. Otros instrumentos, como contadores de centelleo y contadores de cascada, ayudan a identificar las partículas cargadas; también detectan fotones, aunque con una resolución de energía menor que la Bola de Cristal.

Por tanto, el sistema sigue siendo un átomo de positronio, aunque sus partículas constituyentes han intercambiado sus identidades. En algunos estados del positronio, este intercambio no implica nada; mas en otros produce un cambio de signo en la función de onda. Existe un método para saber cómo responde a la conjugación de carga un estado determinado. Cuando se intercambian las cargas, la función de onda del estado queda multiplicada por el número cuántico de conjugación de carga, cuyo valor es +1 para los estados que no se han alterado y −1 para los demás. El número de conjugación de carga, *C*, se conserva en todas las interacciones mediadas por la fuerza electromagnética y la de color.

El fotón, que es su propia antipartícula, tiene un número de conjugación de carga intrínseco de −1. Para un sistema compuesto por varios fotones u otras partículas los valores de *C* de todas ellas se multiplican para dar el número de conjugación de carga total. De ello se sigue que un estado de positronio con *C* igual a +1 debe desintegrarse en número par de fotones; casi siempre dos. Un estado de positronio con *C* igual a −1 dará un número impar de fotones. Como hemos visto antes, la desintegración en un único fotón está prohibida por la conservación de la energía y del momento; así, el mínimo número impar es tres.

¿Cómo se determina el número de conjugación de carga de un estado de positronio? Sumando las magnitudes absolutas del momento angular orbital y del espín total. Si la adición es un número par, *C* valdrá +1; si la suma es impar, *C* será −1. Así, para el estado 2^1P_1, con una unidad de momento angular orbital y con espín total cero, la suma será impar y el valor de *C*, −1. En cambio, los estados 2^3P_0, 2^3P_1 y 2^3P_2 tienen todos una unidad de momento angular orbital y una unidad de espín total. Al ser su suma par, su número de conjugación de carga se cifrará en +1.

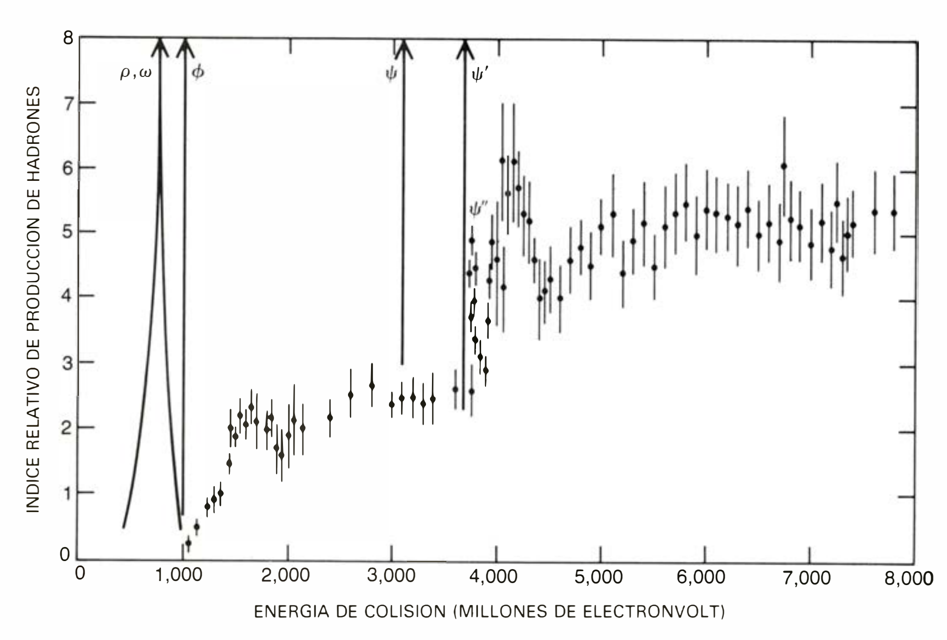
Las distintas propiedades de la función de onda de un estado de positronio influyen sobre la vida media de ese estado. La magnitud de la función de onda correspondiente a una separación entre las partículas igual a cero determina la probabilidad de que éstas colisionen, choque que es condición previa para su aniquilación. Para estados cuyas funciones de onda tengan, a separación nula, la misma intensidad, como el 1^1S_0 y el 1^3S_1, el número de fotones que se emitirán constituye un factor importante que determina la vida media. En efecto, tarda más en emitir más fotones. El estado 1^1S_0, con *C* igual a +1, se desintegra en dos fotones en unos 10^{-10} segundos. El estado 1^3S_1 encierra la misma probabilidad de colisión electrón-positrón, pero al tener *C* igual a −1 debe desintegrarse en tres fotones y su vida media será unas 1000 veces más larga.

Podemos describir el sistema positronio, en su integridad y con gran precisión, a través de la teoría denominada electrodinámica cuántica (QED, en su abreviatura inglesa). Esta teoría atribuye la fuerza entre dos partículas cargadas eléctricamente al intercambio de fotones virtuales emitidos por una de ellas y absorbidos por la otra. Lo de "virtuales" obedece a que los fotones no pueden detectarse directamente en el laboratorio. La probabilidad de que un electrón emita un fotón virtual es proporcional a α, la constante de la ley de Coulomb, cuyo valor se cifra en 1/137.

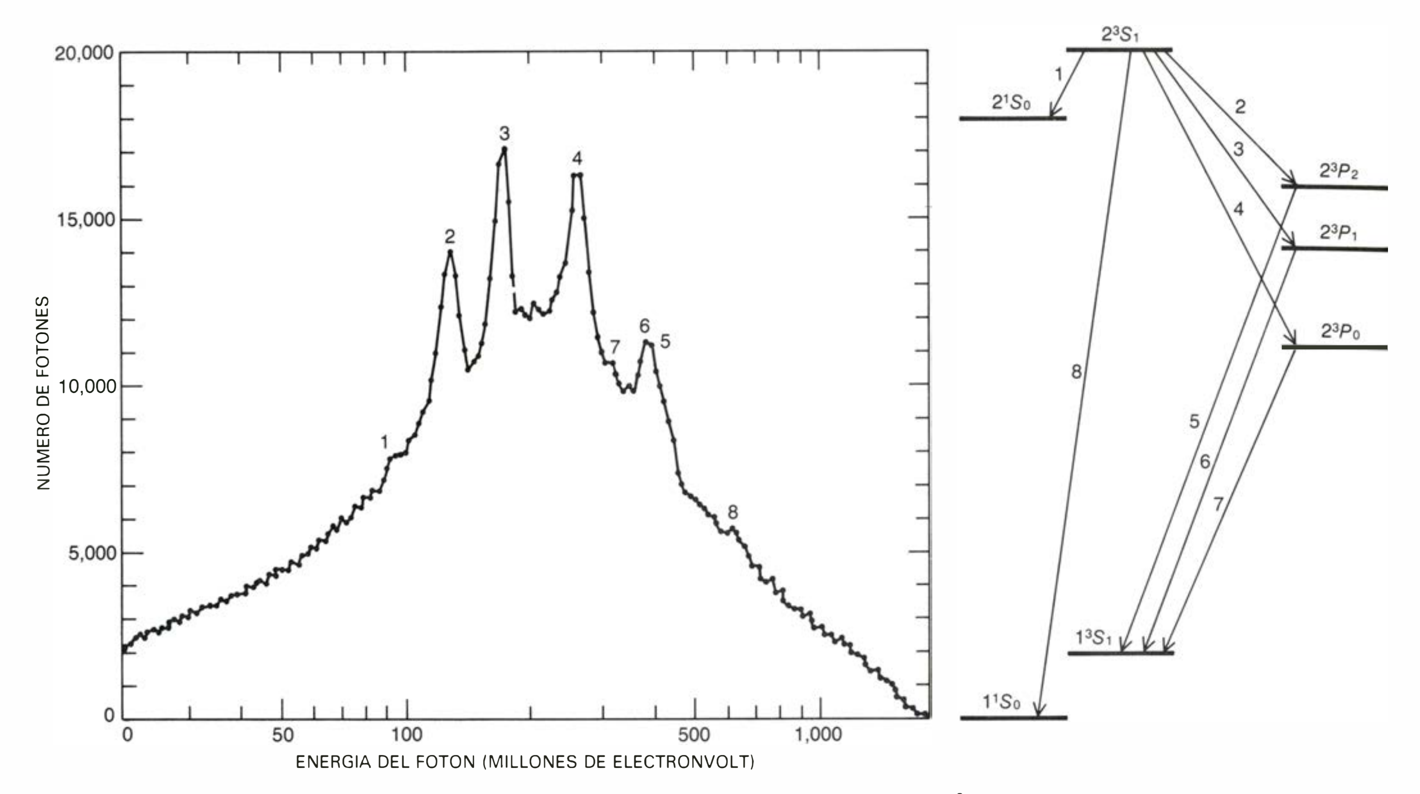
Para calcular la fuerza entre un electrón y un positrón podemos empezar por determinar la fuerza que resulta del intercambio de un simple fotón; resultado que es proporcional a α. Para mayor precisión, incluimos la posibilidad de que sean intercambiados dos fotones; este cálculo añade una pequeña corrección al resultado inicial, proporcional a α^2. El posible intercambio de tres fotones origina una corrección todavía menor, proporcional a α^3. Para llegar al resultado exacto deberíamos realizar una serie infinita de tales cálculos; ahora bien, por ser α pequeño, las correcciones se hacen en seguida despreciables y se obtiene una considerable precisión con sólo incluir los primeros términos de la serie.

En los últimos años se ha desarrollado una teoría de la fuerza de color entre quarks; está inspirada en la QED y se le denomina cromodinámica cuántica (QCD, en su abreviatura inglesa). Así como la QED describe las interacciones de las partículas eléctricamente cargadas, la QCD da cuenta de las fuerzas generadas entre partículas con carga de color. Sin embargo, mientras que existe un único tipo de carga eléctrica, hay tres tipos de cargas de color, usualmente llamadas roja, azul y verde. (Los nombres no tienen nada que ver con lo que en el lenguaje ordinario entendemos por color.) Y, mientras que sólo hay una partícula mediadora de la fuerza electromagnética (el fotón), existen ocho mediadores de la fuerza de color; se les llama gluones.

Tal vez la diferencia más importante entre la QCD y la QED estribe en que los propios gluones acarrean carga de color, al tiempo que el fotón es eléctricamente neutro. Un quark puede emitir y reabsorber gluones continuamente, lo cual produce el efecto de extender sobre una región del espacio su carga de color. Cuando dos quarks o un quark y un antiquark se acercan entre sí suficientemente, las regiones extendidas de carga de color empiezan a sola-



ESTOS PRONUNCIADOS AUMENTOS en la emisión de hadrones (término genérico para las partículas formadas por quarks) denuncian la existencia de sistemas quarkonio. Los datos de energías superiores a 2400 MeV fueron registrados en el SPEAR; los datos de menor energía provienen de anillos de almacenamiento de Francia, Italia y la Unión Soviética. Los picos de la izquierda señalados ρ (rho), ω (omega) y φ representan estados cuasi-ligados de quarks ligeros. El ψ y el ψ' son los estados 1^3S_1 y 2^3S_1 del charmonio. El aumento cerca de 3770 MeV es el estado 3^3D_1 del charmonio, un estado cuasi-ligado designado por ψ''. El aumento registrado en torno a los 4000 MeV se debe a la creación de los mesones encantados *D*.



EXPLORACION DE LAS TRANSICIONES ENTRE estados del sistema charmonio con el detector de la Bola de Cristal. Sólo los estados $^3S_1$ pueden formarse directamente en colisiones electrón-positrón; los demás se presentan cuando uno de los estados $^3S_1$ se desintegra emitiendo un fotón que se detecta con la Bola de Cristal. Los números sobre los picos del espectro de fotones se corresponden con los de las transiciones del diagrama de la derecha.

parse. Debido a ese solapamiento, la fuerza entre las partículas es menor de lo que sería entre dos cargas puntuales separadas por la misma distancia. Al acercarse las partículas entre sí, la fuerza de color entre las mismas, de suyo bastante fuerte, se torna más débil y adquiere la misma forma que la fuerza de Coulomb: su intensidad es inversamente proporcional al cuadrado de la distancia. Esta propiedad de la QCD, descubierta por Politzer, David Gross y Frank Wilczek, los dos últimos de la Universidad de Princeton, apuntaló la esperanza, abrigada por Appelquist y Politzer, de que el sistema charmonio formara estados ligados no relativistas.

No se conoce bien la fuerza que media entre un quark y un antiquark muy alejados, pero hay razones para creer que se trata de una fuerza constante, independiente de la distancia. De ser así, la energía necesaria para separar un quark de un antiquark aumentaría indefinidamente al alejarse uno de otro. Ello explicaría por qué no se han aislado las partículas con carga de color (esto es, quarks y gluones).

Que la fuerza de color de largo alcance no se haya caracterizado todavía en detalle débese al motivo siguiente: la QCD, al contrario que la QED, es una teoría en la que los cálculos resultan engorrosos. La constante α de la QED se sustituye en la QCD por α_s, que es bastante mayor; además, su valor aumenta con la distancia entre las partículas. Por tanto, la probabilidad de que se emita un gluon es grande; pero no es lo bastante pequeña la probabilidad de que dos o tres o más gluones sean emitidos como para que podamos prescindir de ellos. La dificultad que entraña el cálculo de la QCD hace que el quarkonio se convierta en el sistema más importante para el estudio de la fuerza de color. La QCD puede servir de guía para construir modelos sobre esta fuerza, y aquellos someterse luego a prueba experimental en el sistema quarkonio. Los resultados pueden revelar, a su vez, la validez y las propiedades de la teoría.

Para estudiar el quarkonio hay que crearlo, ante todo. ¿Cómo? Mediante la aniquilación de electrones y positrones de alta energía. Desde el punto de vista teórico, el proceso es idéntico a la desintegración del positronio, si bien se produce por un mecanismo algo distinto debido a la mayor energía. El electrón y el positrón, al aniquilarse, originan un simple fotón, cosa que está prohibida en la desintegración del positronio, ya que la energía y el momento no pueden conservarse a la vez. Estas cantidades deben conservarse también en la aniquilación a alta energía, pero el principio de incertidumbre de Werner Heisenberg permite una momentánea violación de la conservación de la energía. Antes de que el fotón haya existido lo suficiente para ser registrado se desintegra en dos o más nuevas partículas. Es un fotón virtual.

En el estado final originado por la desintegración del fotón virtual todas las leyes de conservación tienen que respetarse. Condición que puede cumplirse de forma simple: el fotón virtual origina una partícula y su correspondiente antipartícula. En algunos casos, la pareja partícula-antipartícula consiste en otro electrón y otro positrón, que pueden luego separarse y ser detectados. Pero los productos de la desintegración pueden ser, asimismo, un quark y un antiquark, que no escapan libremente sino que se materializan parejas quark-antiquark adicionales que se ligan al par original, de suerte que lo que se observa en última instancia es siempre un conjunto de hadrones.

Las aniquilaciones a alta enegía tienen lugar en un dispositivo denominado anillo de almacenamiento, por donde circulan haces de electrones y de positrones en direcciones opuestas dentro de una cámara de vacío toroidal. La mayor parte de resultados sobre el charmonio que vamos a describir se lograron con el anillo de almacenamiento SPEAR, que se encuentra en el SLAC. Posee el SPEAR dos regiones de interacción, lo que posibilita el uso simultáneo de dos detectores. Los dos detectores

empleados se distinguen por el nombre familiar que les han impuesto los físicos: Mark II y la Bola de Cristal. Ambos proporcionan información complementaria.

El Mark II lo construyeron ingenieros del SLAC y del Lawrence Berkeley Laboratory de la Universidad de California. Se muestra sobremanera eficaz a la hora de medir la energía de las partículas eléctricamente cargadas, que registra un instrumento denominado cámara de deriva. Hay en ella multitud de finos alambres paralelos desplegados dentro de un volumen cilíndrico lleno de gas. Se aplica un potencial de varios miles de volt a los alambres adyacentes. Cuando una partícula cargada atraviesa la cámara, ioniza el gas y los electrones liberados son entonces atraídos (derivan) por el alambre positivamente cargado más próximo. Cronometrando los pulsos de corriente, puede determinarse, con precisión de unos 0,2 milímetros, la trayectoria de la partícula al atravesar 16 cilindros concéntricos de alambres. La cámara de deriva completa está inmersa en un fuerte campo magnético que obliga a la partícula cargada a seguir una trayectoria curva. A partir del radio de curvatura puede obtenerse el momento de la partícula.

El Mark II contiene instrumentos sensibles a los fotones de alta energía, pero éstos los detecta, con mayor precisión, la Bola de Cristal. Por ser neutro, el fotón no es desviado por un campo magnético, de donde se entiende que la Bola de Cristal no posea campo alguno. El fotón es, en cambio, absorbido en un denso cristal de yoduro sódico originando una cascada de parejas electrón-positrón. La cascada hace que el cristal centellee, esto es, emita muchos fotones de longitud de onda visible, detectados por un tubo fotomultiplicador.

En el diseño y construcción de la Bola de Cristal intervinieron físicos e ingenieros del Cal Tech, Harvard, Princeton y Stanford. El principal componente del detector es un conjunto de 732 cristales simétricamente dispuestos alrededor de la zona de interacción. La energía de un fotón puede determinarse con una precisión de un 2 o 3 por ciento y su dirección con un error de uno o dos grados.

El descubrimiento del mesón ψ en el SLAC debióse al detector Mark I, antecesor del Mark II. Se empleaba entonces el aparato para medir el índice de producción de hadrones en función de la energía del electrón y positrón incidentes. Si los quarks originados por el fotón virtual actuaran inicialmente como partículas libres, el índice de producción de hadrones habría de resultar proporcional a la suma de los cuadrados de las cargas de todos los quarks que podían crearse a una energía dada. En la mayor parte del intervalo de energías explorado, esta hipótesis parecía confirmarse por las mediciones. Sin embargo, cuando se fijó la energía de los haces que colisionaban en el valor preciso de 3095 MeV, el número de hadrones emitidos se centuplicó bruscamente. No tardó mucho en advertirse que este pronunciado aumento delataba la existencia de un nuevo tipo de quark. La excepcional altura y estrechez del pico indicaban también otra cosa. A 3095 MeV, la hipótesis de que el quark y el antiquark actúan como partículas libres falla completamente. El quark y antiquark encantados que constituyen la partícula ψ no nacen libres, sino que están ligados desde el mismo momento de su creación.

El mesón ψ no era el primer ejemplo de un pico pronunciado en el índice de producción hadrónica. Otros tres mesones de corta vida media se habían observado de forma similar a energías más bajas: el ρ (rho) a 776 MeV, el ω (omega) a 782 MeV y el ϕ (phi) a 1020 MeV. Podría pensarse que estos mesones eran estados ligados de los quarks u, d, y s, aunque probablemente fueran estados ligados relativistas. Pero tal explicación sólo en parte es correcta: el ρ, el ω y el ϕ se consideran estados cuasiligados, y no completamente ligados. La distinción radica en que los quarks del hadrón puedan o no aparecer en sus productos de desintegración. Considérese el mesón ϕ, compuesto por un par $s\bar{s}$. Puede desintegrarse por aniquilación mutua; pero éste no es el curso normal de los hechos. Antes bien, el quark s y el antiquark $\bar{s}$ simplemente se separan al tiempo que un nuevo quark u y un antiquark $\bar{u}$ se materializan, formando dos nuevos mesones con la composición $\bar{s}u$ y $s\bar{u}$. Estos mesones se denominan K^+ y K^- y su masa total está 32 MeV por debajo de la masa del ϕ. Este exceso de 32 MeV se convierte en energía cinética de los mesones K.

La partícula ψ podría desintegrarse de forma similar de no ser por un accidente de la naturaleza. El mesón compuesto de un quark c y un antiquark $\bar{u}$ es el D^0; su antipartícula, con la constitución $\bar{c}u$, es el $\bar{D}^0$. La masa de los mesones neutros D es de 1865 MeV; por tanto, la masa de la pareja partícula-antipartícula será de 3728 MeV, 633 MeV por encima de la masa de la ψ. Así, el mesón ψ no puede desintegrarse en dos mesones D porque la desintegración violaría la conservación de la energía. El c y el $\bar{c}$ no pueden alejarse uno de otro, sino que se desintegran por aniquilación mutua formando gluones, que luego se transforman en

LA FUERZA ENTRE QUARKS, denominada fuerza de color, parece obedecer una ley significativamente distinta de la que describe la fuerza electromagnética. La intensidad de la fuerza electromagnética entre dos partículas varía en proporción inversa al cuadrado de la distancia que las separa; tal ley de fuerza corresponde a una línea recta inclinada en este gráfico logarítmico. A distancias extremadamente cortas, la fuerza de color también parece seguir una ley de inverso al cuadrado; ahora bien, más allá de 10^{-13} centímetros la fuerza puede tener un valor constante e independiente de la distancia. Si la fuerza de largo alcance es constante, los datos del quarkonio sitúan su magnitud en torno a las 16 toneladas. La ley de fuerza mostrada se basa en un modelo desarrollado por John Richardson, cuando estaba en el SLAC.

quarks y posteriormente en hadrones. Este proceso es bastante más lento y hace que la vida media del mesón ψ dure cientos de veces más que lo que, de otra forma, debería ser.

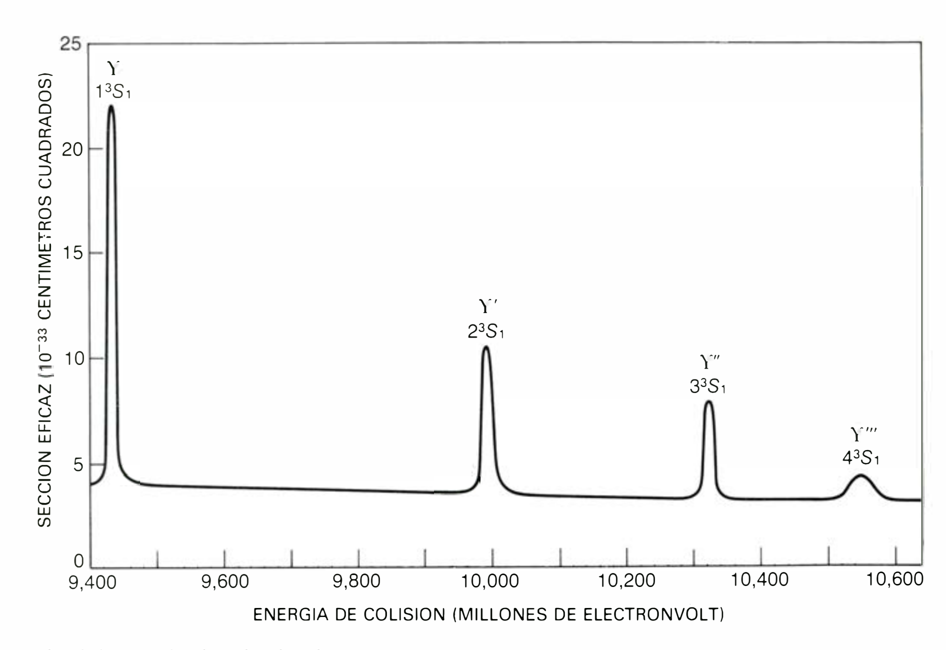
En analogía con el estado de menor energía del positronio, se supuso que el ψ tenía un momento angular orbital igual a cero. Sin embargo, como se crea a partir de un fotón virtual, su momento angular total debe ser igual al del fotón, esto es, una unidad. Por tanto, el ψ constituye el estado 1^3S_1 del charmonio, es decir, el estado con n igual a 1, una unidad de espín total y carente de momento angular orbital.

Inmediatamente después del descubrimiento del ψ se sugirió que el estado 2^3S_1, con el mismo espín total y momento angular orbital n igual a 2, podría también ser un estado ligado. Se inició su búsqueda aumentando la energía de los anillos de almacenamiento en incrementos de 2 MeV. Dos semanas más tarde, se descubrió la partícula denominada ψ', a una energía de 3684 MeV, que se interpreta ahora como el estado 2^3S_1 del sistema $c\bar{c}$. Si su energía fuese 44 MeV más alta, podría desintegrarse en mesones D y sería sólo cuasiligada.

Si el modelo del charmonio es correcto, tiene que haber por lo menos otros cinco estados ligados del charmonio que podrían detectarse en las desintegraciones del ψ y del ψ'. Dos de estos estados son los estados S singletes que únicamente se diferencian del ψ y del ψ' en que su espín total, así como su momento angular orbital, son cero. Se designan por 1^1S_0 y 2^1S_0 y se les denomina mesones η (eta) encantados. Más concretamente, se les llama η_c y η'_c en analogía con el mesón η que es un estado 1^1S_0 de quarks ligeros. Se esperaba que las masas del η_c y del η'_c estuvieran unos 100 MeV por debajo de las del ψ y del ψ', respectivamente.

Los otros tres estados esperados eran los tripletes cuyo momento angular orbital valiera uno: 2^3P_0, 2^3P_1 y 2^3P_2. Fueron denominados mesones χ (ji) por los experimentadores que los buscaban. Se les asignó una masa unos 220 MeV por debajo de la masa del ψ' y una distancia mutua cifrada entre 50 y 150 MeV. Hay un sexto estado predicho por el modelo del charmonio; se trata del estado singlete con espín total cero y momento angular orbital uno, es decir, el 2^1P_1. Sin embargo, se espera una baja probabilidad de conversión entre este estado y los demás, lo que dificulta grandemente su detección y, en efecto, no se ha encontrado todavía.

Hay dos razones que justifican la imposibilidad de crear directamente los estados S singletes y los estados P tripletes del charmonio en aniquilaciones electrón-positrón. En primer lugar, aparte de la configuración 2^3P_1, los demás no tienen el mismo momento angular total que el fotón virtual y, por tanto, su producción directa violaría la conservación del momento angular. En segundo lugar, en los cinco estados la suma de las magnitudes absolutas de su espín total y de su momento angular orbital es un número par (cero para los estados S y 2 para los estados P). Para todos los estados C es, por tanto, +1, mientras que el valor para el fotón virtual es −1. En cambio, para el ψ y el ψ' el valor de C es −1 y, en consecuencia, pueden formarse directamente. Los estados en que C vale +1 pueden formarse únicamente cuando el ψ o el ψ' se desintegran en un fotón y una configuración del charmonio de menor energía, proceso que respeta todas las leyes de conservación.



POR SISTEMAS BOTTOMONIO se conocen los formados por un quark "bottom" ("fondo") y un antiquark "bottom". Más pesados que los correspondientes estados del charmonio, pueden revelar con mayor claridad la naturaleza de las fuerzas entre quarks. Mientras dos de los estados $^3S_1$ del charmonio (el ψ y el ψ') son sistemas totalmente ligados, existen tres estados ligados análogos de bottomonio. Se designan por la letra griega Υ (hypsilon), Υ' y Υ''. Un cuarto sistema bottomonio, el Υ''', es cuasi-ligado. El gráfico se basa en datos obtenidos en el anillo de almacenamiento CESR de la Universidad de Cornell. La altura de los picos aparece en términos de la sección eficaz, que mide la probabilidad de crear una partícula.

La primera prueba sobre la existencia de los mesones χ se obtuvo en un experimento realizado en el DORIS, un anillo de almacenamiento similar al SPEAR, que se encuentra en el Deutsches Elektronen-Synchrotron (DESY) en Hamburgo. Otros experimentos posteriores, basados en diferentes métodos de detección, han identificado también a los mesones χ. Sus masas están 135, 180 y 270 MeV por debajo de la del ψ', lo que confirma las predicciones del modelo del charmonio.

En SPEAR, un grupo de físicos de las universidades de Maryland, Pavía (Italia), Princeton, California en San Diego y Stanford midieron las masas de los estados χ y la proporción en la que éstos aparecían como producto de desintegración del ψ'. Utilizaron un primitivo detector basado en cristales de yoduro sódico con el que no detectaban directamente las partículas χ, sino el fotón emitido en la transición del ψ' al χ. La energía de este fotón está claramente definida y viene a ser igual a la diferencia de masa entre los estados. Se recurrió al Mark I para sacar mediciones, buscando sucesos en los que el χ se desintegrara en varias partículas cargadas. Un χ se desintegra ocasionalmente en dos piones positivos y dos negativos; a partir de la energía y de la distribución angular de los piones cabe deducir la masa del estado χ. Los grupos de la Bola de Cristal y del Mark II acaban de realizar mediciones de más alta resolución.

¿Cómo determinar los números cuánticos de momento angular de los mesones χ? A partir de la naturaleza de los hadrones emitidos en su desintegración y de la distribución angular de los fotones. Los primeros experimentos mostraron que ninguna de las tres partículas podían ser los estados $^1S_0$, es decir, el η_c o el η'_c. Además, si se suponía que el momento angular total de los tres mesones era 2, 1 y cero respectivamente, sólo había una forma de compaginar los momentos angulares con las masas observadas. El estado que poseyera una mayor masa debía ser el de

momento angular 2, el siguiente debía tener momento angular 1 y el más ligero sería el de momento angular cero, formando, de este modo, la secuencia de estados 2^3P del charmonio. Un estudio más detallado hecho posteriormente con la Bola de Cristal confirmó estas asignaciones de momento angular.

La historia de los mesones η_c y η'_c no recorrió un camino tan recto. Los experimentalistas que trabajaban en el DORIS anunciaron la prueba de un candidato a η_c en los primeros años de exploración del espectro del charmonio. Observaron un número anormalmente grande de desintegraciones del mesón ψ en tres fotones, en las que dos de los fotones tenían una masa conjunta de 2830 MeV; esta masa es inferior a la del ψ en 265 MeV. Interpretaron tales sucesos como si se tratara de la emisión de un fotón de energía bastante baja por parte del ψ que, con ello, se transformaba en un estado η_c. Este último empleaba, entonces, su masa de 2830 MeV en formar los dos fotones adicionales observados. La señal parecía impresionante, pero había un importante fondo de sucesos espurios y ningún otro experimento pudo proporcionar una confirmación convincente de los hallazgos del DORIS.

Hubo también otro supuesto candidato a η_c: se fundaba en cuatro sucesos registrados con el detector Mark II y en otros tantos de los experimentos del DORIS. En estos sucesos, el mesón ψ' emitía un fotón y se convertía en el mesón η'_c, el cual emitía, entonces, un segundo fotón al caer al nivel del ψ. La masa de este candidato a η'_c era de 3454 MeV, es decir, 230 MeV por debajo de la masa del ψ'.

Ninguno de los candidatos se ajustaba fácilmente al modelo del charmonio. Las masas eran demasiado bajas. La esperada para el η_c fue calculada en 1978 por Michael Shifman, Arkady Vainshtein, Michael Voloshin y Valentin Zakharov, del Instituto de Física Teórica y Experimental de Moscú. A partir de principios teóricos basados en la QCD demostraron que el mesón η_c tendría una masa inferior en unos 100 MeV a la masa del ψ, dentro de un error de más o menos 20 o 30 MeV. Sobre esta base concluyeron sin ambigüedad que o bien el candidato a η_c (con una masa 265 MeV por debajo de la del ψ) no era el mesón η_c, o bien el modelo simple del charmonio no valía.

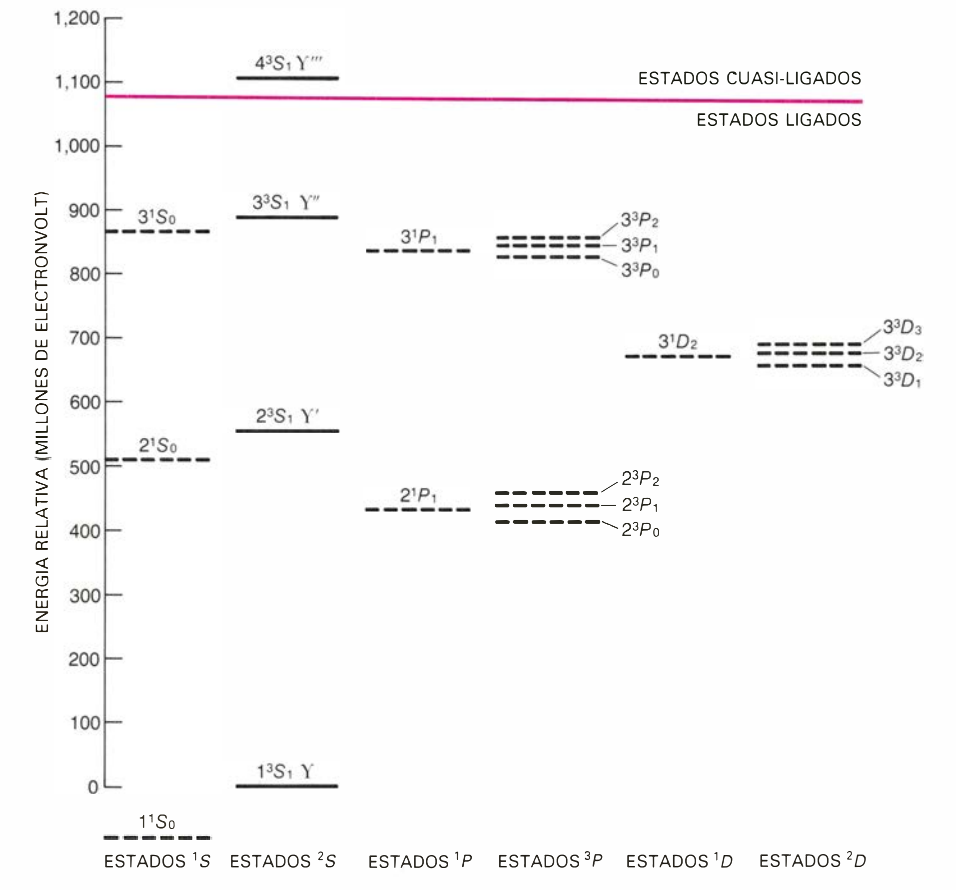
Llegados a este punto, la situación de la hipótesis del charmonio se tornó insegura. Se había encontrado los tres mesones χ, cuyas propiedades estaban en perfecto acuerdo con las predicciones del modelo, en tanto que los candidatos a estados η_n se alejaban del modelo por razón de su masa y de otras ciertas propiedades. Pero no estaba claro si el modelo se entendía lo suficientemente bien hasta el punto de confiar en él allí donde parecía contradecir los datos experimentales.

En 1978, los experimentalistas que trabajaban con la Bola de Cristal empezaron su cosecha de datos. Al año siguiente, repitieron las medidas de la desintegración del mesón ψ en tres fotones, el único modo de desintegración en el que se había observado el candidato a η_c. La Bola de Cristal, capacitada para detectar muchas más desintegraciones de este tipo que las que se habían registrado con los aparatos del DORIS, podía resolver la energía de los fotones con mayor precisión. Mas el candidato no apareció. El anterior experimento fue, casi con toda certeza, erróneo, aunque la razón no está todavía clara. Quizás hubo una fluctuación estadística en el número de sucesos de fondo o una mala interpretación de la respuesta del detector a los mismos.

El candidato a mesón η'_c tuvo un final semejante. En los experimentos del DORIS y en los del SPEAR, con los detectores Mark II y la Bola de Cristal, se rastreó la zona de energía correspondiente al candidato anunciado. Aunque los nuevos experimentos eran más finos que los anteriores, no encontraron prueba alguna de la existencia del mesón η'_c. No había tal partícula de 3454 MeV de masa.

Por fin, semejante estado de incertidumbre acabó en las postrimerías de 1979 con el descubrimiento de la partícula que parece ser el verdadero η_c por parte del equipo que trabaja con la Bola de Cristal. El nuevo candidato a η_c tiene una masa de 2980 MeV, es decir, unos 115 MeV menos que el ψ, en concordancia con las predicciones del modelo del charmonio. Este estado de 2980 MeV se observó, por vez primera, como un pico, pequeño y significativo, en el espectro de fotones emitido en la desintegración del ψ'. Posteriormente, se identificó un pico de la misma energía en el espectro del fotón emitido en la desintegración del ψ. La existencia del η_c recibió confirmación inmediata con el detector Mark II, que pudo registrar varias de sus desintegraciones en hadrones.

En agosto de 1981, se descubría, muy cerca de donde se había predicho, el último de los cinco estados del charmonio con C igual a +1. Quienes trabajaban con la Bola de Cristal anunciaron la ob-



NIVELES DE ENERGIA DEL BOTTOMONIO. Se espera que incluyan más estados ligados que los del charmonio. Además de los estados ligados S y P, pueden detectarse ciertos estados ligados D, con dos unidades de momento angular orbital. De momento, solamente se han observado los estados $^3S_1$.

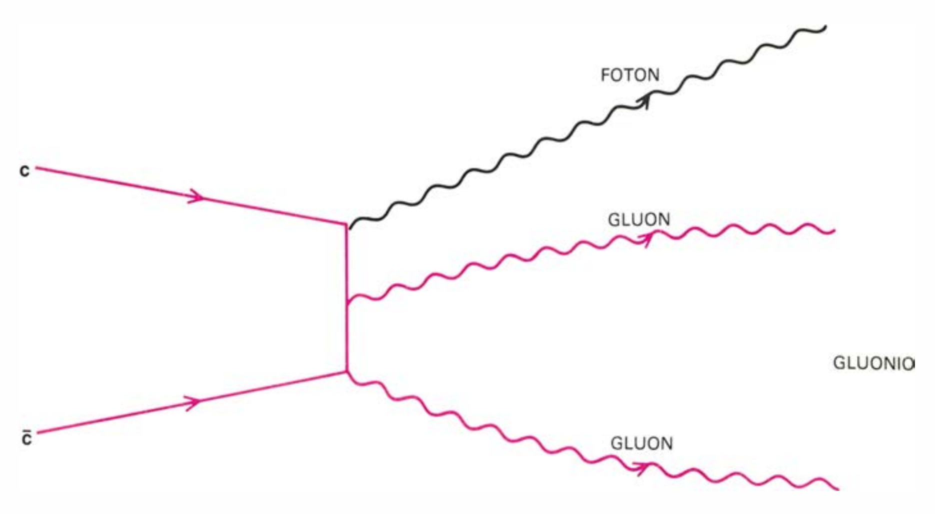
servación de otro pequeño pico en el espectro de los fotones emitidos al desintegrarse los mesones ψ'. El pico se hallaba a una energía de 3592 MeV, unos 92 MeV por debajo de la masa del ψ', cayendo perfectamente en la región de masas predicha para el η'_c. Con este descubrimiento, los enigmas del η_c y η'_c, parecen resueltos y tenemos vía libre para aplicar el modelo del quarkonio al estudio de la fuerza de color.

Probablemente el mejor sistema quarkonio que dispongan hoy los experimentalistas sea el sistema $b\bar{b}$ o "bottomonio". Debiera aventajar al charmonio, ya que posee una masa bastante mayor y se aproxima más, por tanto, a un verdadero sistema no relativista.

Cuando en 1977 se descubrió el Y en el Fermilab, se le interpretó como el estado 1^3S_1 del bottomonio, en completa analogía con el ψ, que es el mismo estado del charmonio. El descubrimiento iba acompañado del anuncio de la probable observación del estado 2^3S_1, denominado Y' y de la observación provisional del estado 3^3S_1, el Y''. La existencia de los tres estados ligados del bottomonio (en comparación con los dos del charmonio) constituía una importante predicción del modelo propuesto por Eichten y Gottfried.

Debido a la mayor energía de ligadura del bottomonio, el conjunto de estados ligados asociados tenía que ser aún más rico que el del charmonio. Por cada partícula Y había de existir un estado *S* singlete algo menos masivo: son los estados 1^1S_0, 2^1S_0 y 3^1S_0, denominados η_b, η_b' y η_b''. Análogamente deberían existir tres partículas χ_b, para representar el triplete de estados 2^3P, esto es, 2^3P_0, 2^3P_1 y 2^3P_2. Sin embargo, en el sistema bottomonio los propios estados $3P$ deberían ser ligados, al contrario de lo que ocurre en el charmonio. Además, los estados *D* de menor energía, con dos unidades de momento angular orbital, también tendrían que existir como estados ligados. Corresponden a las configuraciones 3^3D_1, 3^3D_2 y 3^3D_3.

Aparte de las partículas Y, no se ha observado todavía ningún otro sistema bottomonio. Los experimentos encierran una dificultad mayor que los del charmonio, cuya razón obedece a que los sistemas $b\bar{b}$ se producen en menor proporción que los sistemas $c\bar{c}$. En varios laboratorios se está llevando a cabo una cuidadosa búsqueda. El anillo de almacenamiento electrón-positrón CESR, en Cornell, tiene casi el rango de energías óptimo para crear sistemas bottomonio; y otros anillos de almacenamiento, capaces de alcanzar energías aún más altas, se han construido en el SLAC y en el DESY. En ese orden de cosas, notemos que se ha desmontado del SPEAR la Bola de Cristal para instalarla en una región de interacción del DESY, un anillo más energético. En la próxima generación de experimentos debieran observarse ya los estados χ_b e incluso tal vez los η_b.



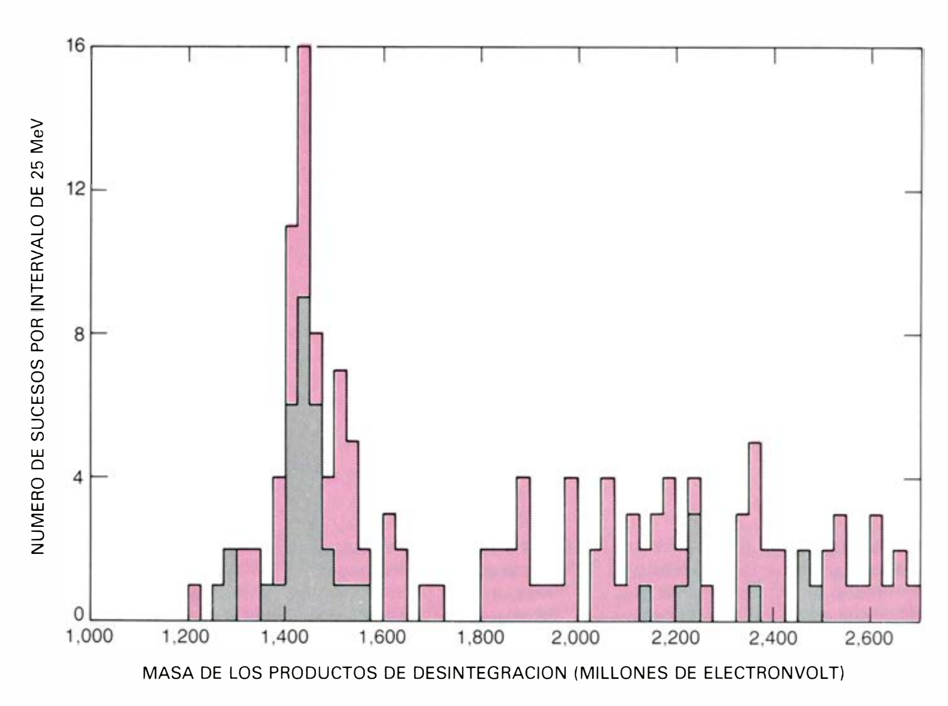
ES EL GLUONIO un hipotético estado ligado y formado enteramente por gluones, partículas que ordinariamente actúan como intermediarios de la fuerza de color entre los quarks. Por acarrear cargas de color, los propios gluones están sometidos a la fuerza de color y pueden formar estados ligados. Si el gluonio realmente existe, es probable que se cree en sucesos que producen abundantes gluones. Un mecanismo plausible es el de la desintegración de un mesón ψ en un fotón y dos gluones; en tales circunstancias, los gluones tendrían necesariamente cargas de color opuestas y podrían constituir así un estado ligado.

¿Qué enseñanzas hemos sacado del quarkonio que nos revelen mejor la naturaleza de la fuerza entre los quarks? Puesto que los distintos estados del quarkonio difieren en la separación media entre el quark y el antiquark, las energías de estos sistemas nos informan sobre la magnitud de la fuerza en cierto rango de distancias. Los experimentos realizados hasta hoy miden la fuerza de color a distancias que van desde 10^{-13} centímetros, que viene a suponer el tamaño de un hadrón ordinario, hasta 2×10^{-14} centímetros, una quinta parte. A partir de las medidas podemos construir modelos sobre cómo varía la fuerza con la distancia entre los quarks.

La QCD sugiere el siguiente comportamiento de la fuerza de color: varía probablemente como el inverso del cuadrado de la distancia cuando los quarks están muy juntos, mientras que se hace constante cuando los quarks están separados. Una forma plausible de aproximar dicha fuerza es suponer que se trata de la suma de una fuerza tipo inverso-cuadrado y una fuerza constante. Esta ley de fuerza modelo puede expresarse matemáticamente por la ecuación $F = a/r^2 + b$, donde *F* simboliza la fuerza, *r* la distancia entre los quarks y *a* y *b* son constantes a determinar experimentalmente. A distancias muy pequeñas, la razón a/r^2 es grande y constituye la principal contribución a la fuerza total. A grandes distancias, en cambio, el término a/r^2 es despreciable, y la fuerza se iguala, grosso modo, a *b*.

¿Cómo comprobar esta ley de fuerza modelo? Intentando ajustar sus predicciones a los datos del quarkonio. Si la curva resultante tiene la forma correcta, puede preguntarse qué valores de *a* y *b* dan el mejor ajuste. Resulta, efectivamente, que la forma de la curva es coherente con los datos y que el valor de *b* es de unas 16 toneladas. En otras palabras, dos quarks se atraen con una fuerza de por lo menos 16 toneladas con independencia de cuán alejados estén uno de otro. Con esto por delante, se hace más fácil comprender cómo es que nunca se haya podido extraer de un hadrón un quark. Adviértase que, con este razonamiento, no hemos probado que exista una fuerza constante entre los quarks; difícilmente puede probarse cuando los datos sólo se extienden a distancias del orden de 10^{-13} centímetros. Lo que hemos mostrado es que los datos disponibles se muestran coherentes con una fuerza de largo alcance constante y que, si existe, su magnitud sería de unas 16 toneladas.

Más difícil resulta especificar el valor de *a* en la ecuación de la ley de fuerza. La razón estriba en que *a* no es una constante, sino que depende en cierta medida de *r*. El valor de *a* está relacionado con la probabilidad de que un quark emita un gluón, cantidad que hemos designado por α_s. Al contrario que α (probabilidad de que un electrón emita un fotón), α_s varía con la distancia, haciéndose más pequeña a medida que el quark y el antiquark se acercan entre sí. En el rango de distancias explorado con el sistema quarkonio, el



UNA POSIBLE PRUEBA DEL GLUONIO quizá fuera la partícula ι (iota), descubierta el año pasado en el SPEAR. La ι, que tiene una masa de 1420 MeV y momento angular nulo, no encaja directamente en ninguna de las familias de mesones quark-antiquark establecidas. Podría tratarse de un estado ligado de dos gluones, aunque hay otras explicaciones alternativas. Aquí se han seleccionado desintegraciones del mesón ψ registradas por el detector Mark II, incluyendo sólo sucesos que contenían cierto conjunto de tres partículas entre los productos de la desintegración. Se asignó cada suceso a un intervalo de 25 MeV, de acuerdo con la masa conjunta de las partículas. El pico entre 1400 y 1450 MeV está causado por la ι. Las barras grises representan un subconjunto de datos seleccionados con criterios más restrictivos.

valor de α_s parece ser aproximadamente 1/5, que merece la pena contrastarse con el de α, cifrado en 1/137.

El interés que encierra el quarkonio deriva en gran parte del estrecho paralelismo que existe entre él y el positronio y entre la QCD y la QED. Pero en sistemas gobernados por la fuerza de color se presenta un fenómeno que no tiene equivalente en sistemas electromagnéticos. Se trata de una partícula compuesta exclusivamente de gluones ligados por sus cargas de color. El análogo electromagnético sería un estado ligado de dos fotones; tal sistema no puede existir, empero, ya que el fotón carece de carga.

Para referirse a los hipotéticos estados ligados gluónicos se suele utilizar el término "glueballs" (bolas gluónicas). Aquí los llamaremos sistemas "gluonio", al objeto de destacar su similitud con el quarkonio. Si existen, uno de los procesos donde se crearán con mayor probabilidad será en la desintegración de un quarkonio pesado. De la misma forma que el estado 1^3S_1 del positronio se desintegra en tres fotones, el estado 1^3S_1 del charmonio (esto es, la partícula ψ) produce generalmente tres gluones. Estos, a su vez, crean pares quark-antiquark, que se combinan y originan los hadrones. Mas en un 10 por ciento de las ocasiones, el charmonio se desintegra en un fotón y dos gluones; precisamente en estos casos, el gluonio podría tener una mayor probabilidad de formarse. Los dos gluones emitidos en la desintegración tienen cargas de color opuestas y podrían formar directamente un estado ligado neutro en color.

No está del todo claro cómo reconocer el gluonio en la hipótesis de que se cree. Cabe incluso que una partícula vacile continuamente entre estados quark-antiquark y estados gluónicos, en cuyo caso el concepto de gluonio no tendría mucho sentido. Si el gluonio tiene una identidad fija, una de las características que lo distinguirían sería el espectro de estados de momento angular. Por ser el espín de los gluones 1 en vez de 1/2, se espera que el estado gluónico de menor energía sea un estado S con un momento angular total igual a cero o 2, lo que contrasta con el quarkonio, dotado de un momento angular total nulo o de valor unidad. Otro indicio lo constituye lo siguiente: el gluonio debería producirse en procesos ricos en gluones (como la desintegración de un quarkonio en un fotón y dos gluones) con una abundancia mayor que en otras interacciones de hadrones. El gluonio se manifestaría también ante la imposibilidad de encajarlo en ninguna de las familias de mesones establecidas.

Hace un par de años, Daniel Scharre, del SLAC, descubrió en los datos del detector Mark II señales de un curioso modo de desintegración de la partícula ψ. En aproximadamente tres desintegraciones de cada 1000, el ψ daba un fotón y un mesón cuya masa era de unos 1420 MeV. Scharre lo identificó como el mesón E, que es uno de los estados 2^3P_1 de quarks ligeros y cuyo momento angular total es, en consecuencia, uno. El descubrimiento constituyó una sorpresa, pues se trataba de un mesón bastante raro, y aunque un tres por 1000 era una proporción pequeña entre todas las desintegraciones del ψ, se considera un alto índice de producción para el mesón E.

Michael Chanowitz y otros sugirieron que podía tratarse de un caso de error de identificación. La partícula observada en la desintegración del ψ se había identificado a partir de su masa y su modo de desintegración más prominente, que eran los mismos que los del mesón E. Chanowitz notó que la nueva partícula difería del E en ciertos detalles de su producción y desintegración. Y afirmó que lo visto no era el mesón E, sino una nueva partícula cuyo momento angular total era nulo.

En agosto del año pasado se confirmó que Chanowitz andaba acertado en su razonamiento: el equipo de la Bola de Cristal anunció la medida del momento angular de la partícula de 1420 MeV, y resultó ser cero. Como no podía tratarse del mesón E, se le dio un nuevo nombre: ι (iota). Al propio tiempo, el equipo de la Bola de Cristal comunicaba el descubrimiento de otra nueva partícula entre los productos de desintegración del ψ. Al igual que la ι, esta nueva partícula aparece conjuntamente con un fotón, pero su masa es mayor: 1640 MeV. A esta segunda partícula se le ha dado el nombre θ (theta) y se ha observado que posee dos unidades de momento angular total.

¿Qué son las partículas ι y θ? No hay contestación fácil. Si corresponden a estados quark-antiquark, la ι debe pertenecer a la familia $^1S_0$ de mesones y el θ a la familia $^3P_2$. No puede tratarse de los estados $n=1$ de estas familias, porque éstos se encuentran ya ocupados por partículas conocidas. Podrían ser los estados $n = 2$. Quizá mesones "dobles" exóticos compuestos por dos quarks y dos antiquarks. Y, por supuesto, podría tratarse de estados de gluonio. Cada una de estas posibilidades implica un distinto espectro de nuevas partículas acompañantes. Si los experimentos demuestran que el espectro es el característico del gluonio, los mesones ι y θ se convertirán en los primeros ejemplos de una nueva forma de materia: partículas con la fuerza que media entre los quarks, pero sin quarks.

Imágenes por RMN en medicina

La resonancia magnética nuclear, RMN, revela la distribución de los átomos en una muestra de material. Si se aplica al cuerpo, genera imágenes de la estructura interna sin necesidad de emplear rayos X

Ian L. Pykett

Desde hace muchos años se viene visualizando, con fines médicos, el interior del cuerpo humano mediante el registro de su absorción diferencial de los rayos X. Ese método convencional de radiografía presenta, sin embargo, una importante limitación: es incapaz de distinguir entre estructuras superpuestas. Barrera que acaba de saltarse con el desarrollo de la tomografía computarizada de rayos X, o TC; esta técnica reconstruye matemáticamente los datos radiográficos, tomados desde varias direcciones, para reproducir secciones transversales de cualquier parte del cuerpo que se desee. Aunque el procedimiento se ha convertido en una poderosa herramienta de diagnóstico, la información que proporcionan sus imágenes es, fundamentalmente, de tipo anatómico; poco dicen del estado funcional o fisiológico de los órganos internos. Es más, las propiedades de absorción de rayos X de algunas lesiones son muy similares a las de los tejidos colindantes, hasta el punto de escaparse a la detección en los exámenes TC cuando no llegan a alterar el tamaño y la forma del órgano. Por otro lado, los rayos X comportan, incluso a pequeñas dosis, un riesgo claro de inducción de lesiones fisiológicas.

Una nueva técnica destinada a la obtención de secciones transversales del cuerpo humano, que no expone al paciente a la acción de radiaciones ionizantes, está llamando a la puerta de la aplicación clínica: nos referimos a la formación de imágenes por resonancia magnética nuclear. Las imágenes construidas por RMN, además de ofrecernos una información anatómica equiparable a la que nos suministra el barrido por TC, permitirán distinguir, de un modo mucho más fino, entre el tejido sano y el enfermo. Esperanza de futuro que tiene su fundamento en la ya demostrada capacidad de la espectroscopía por RMN a la hora de desvelar la intrincada conformación de las moléculas orgánicas y llevarnos a una visión más profunda de la dinámica implicada en los procesos químicos. A lo largo de varios años, los bioquímicos se han servido de las técnicas de RMN para investigar y seguir las reacciones metabólicas en animales de laboratorio y en humanos. Mas el desarrollo experimentado hoy en los métodos para la representación plástica de la información recabada por la RMN le proporciona al clínico un nuevo medio de diagnóstico muy valioso.

Felix Bloch, de la Universidad de Stanford, y Edward M. Purcell, de Harvard, establecieron las bases experimentales de la espectroscopía de RMN hace ya más de treinta años. Por cuyo trabajo recibieron el Nobel en 1952. Pero ya desde los años 20 de nuestro siglo se sabía que muchos núcleos atómicos tenían un momento angular derivado de su propiedad intrínseca de rotación, el espín. Por estar los núcleos eléctricamente cargados, el espín corresponde a un flujo de corriente en torno al eje de espín, que genera a su vez un pequeño campo magnético. Así pues, el núcleo de espín no nulo poseerá un momento magnético, un dipolo, asociado al mismo. Sólo los núcleos con un número impar de nucleones (protones o neutrones) tienen un espín total o neto y se prestan, por consiguiente, a la espectroscopía por RMN.

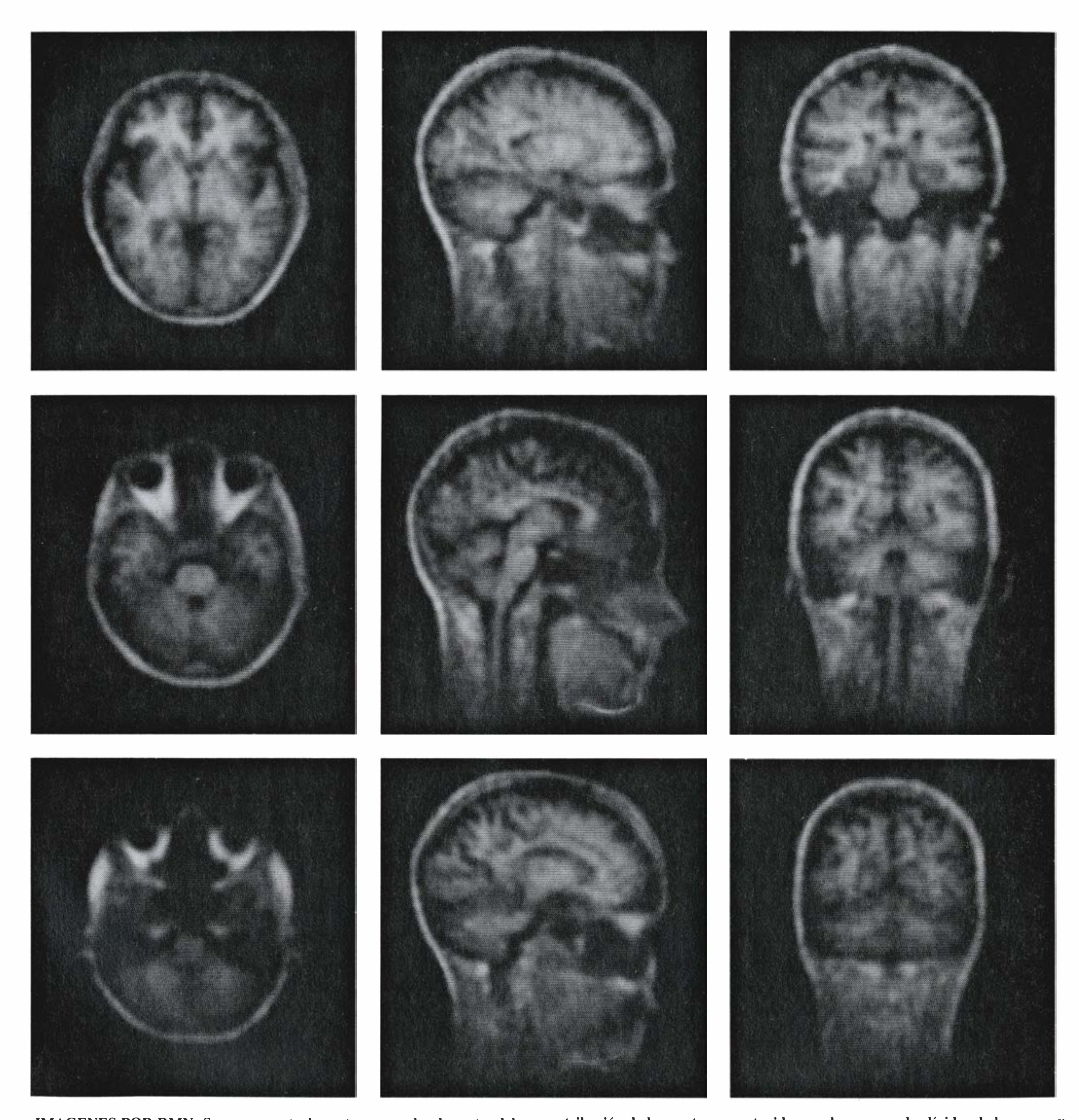
En general, los dipolos magnéticos de los núcleos con espín apuntan en direcciones aleatorias. Pero si los colocamos en un campo magnético, se orientarán según las líneas de inducción del campo, o líneas de fuerza. Para los núcleos de espín 1/2, como los protones (núcleos de hidrógeno, ^{1}H), las únicas orientaciones posibles de los momentos magnéticos son las del sentido paralelo al campo o antiparalelo al mismo (en sentido opuesto). Las dos orientaciones tienen energías ligeramente diferentes y constituyen lo que se describe como desdoblamiento de Zeeman de los niveles de energía. En el caso de los protones la diferencia entre el número de protones con espín "hacia arriba" –up– (paralelo) y espín "hacia abajo" –down– (antiparalelo) es muy pequeña: solamente una parte por 10^8, con un ligero exceso en el estado de menor energía (espín hacia arriba).

¿Cómo predecir el comportamiento magnético de la totalidad de la población de los núcleos? Mediante la definición de un vector macroscópico de magnetización general, M, que representa el efecto de todos los momentos magnéticos de los núcleos de una determinada especie nuclear presente en la muestra del material sometido a examen. En ausencia de un campo magnético externo, la magnetización global, evidentemente, vale cero. Sin embargo, al someter la muestra a un campo magnético, los dipolos nucleares se orientan y producen una magnetización finita en equilibrio, que apunta en dirección paralela al campo magnético aplicado. Esta dirección define, por convención, el eje z.

Los núcleos en rotación se comportan de forma muy similar a pequeños giroscopios. Si se inclina el eje de un giroscopio en rotación alejándolo de la vertical, el giroscopio girará alrededor de su anterior eje en un movimiento que describe la pared de un cono. Este movimiento recibe el nombre de precesión. Análogamente, si la magnetización global M, correspondiente a una agrupación de núcleos en rotación en un campo magnético, se aleja de la dirección z, M efectuará un movimiento de precesión alrededor del eje z. Esta inclinación se consigue aplicando un campo magnético mucho menor que gire en el plano x-y, en ángulo recto con el campo estático (sin rotación). En la práctica, se aplica el campo magnético en rotación rodeando la muestra con una bobina conectada a una fuente de radiofrecuencia. Para inclinar el vector del espín macroscópico alejándolo del eje z, la frecuencia de la radiación electromagnética aplicada debe ser igual a

la frecuencia de precesión natural de los núcleos de la muestra, de ahí la expresión resonancia magnética nuclear.

Una simple relación matemática conjuga la frecuencia de resonancia, a menudo denominada frecuencia Larmor, con el valor del campo magnético estático aplicado externamente. La frecuencia equivale a la fuerza del campo multiplicada por la "constante magnetogírica", valor único para cada especie nuclear de espín no nulo. Para los núcleos de hidrógeno (protones) en un campo magnético de un tesla (10.000 gauss) la frecuencia de resonancia es de 42,57 megahertz (MHz), o 42,57 millones de ciclos por segundo. Para los núcleos del isótopo 31 del fósforo (^{31}P), en el mismo campo, la frecuencia de resonancia es de 17,24 MHz; para los núcleos del sodio ^{23}Na es de 11,26 MHz.



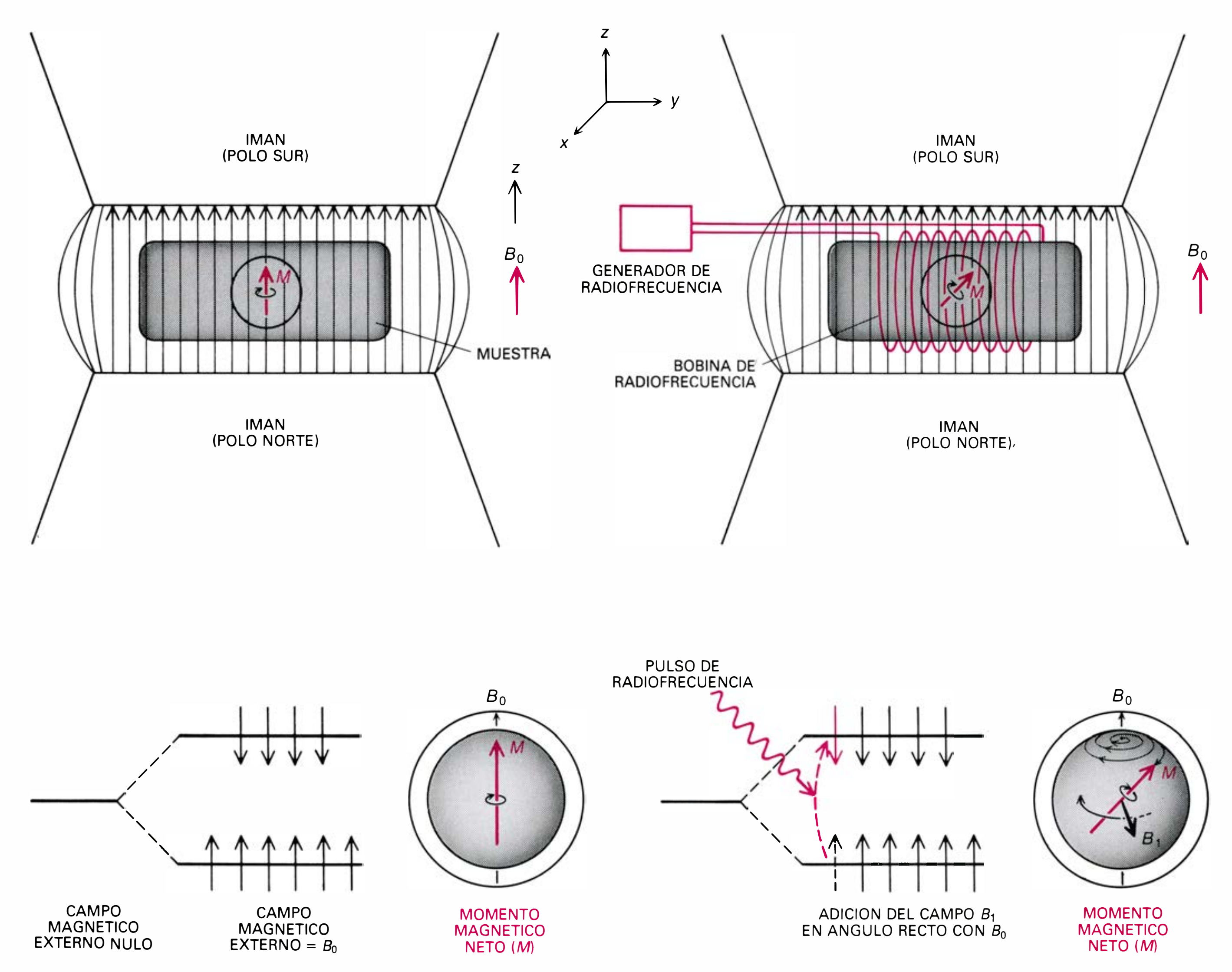
IMAGENES POR RMN. Son representaciones transversales de cortes delgados del cuerpo, obtenidas aplicando ondas de radio a núcleos atómicos susceptibles, que se orientan con precisión en un campo magnético. Son núcleos susceptibles aquellos que tienen un número impar de nucleones (protones o neutrones) y que exhiben, por tanto, un espín neto. Si se excitan con ondas de radio los núcleos del hidrógeno (protones), los de más amplia distribución en los seres vivos, revelan su localización emitiendo durante un breve período de tiempo una señal a una frecuencia precisa. A partir de estas señales pueden reconstruirse las imágenes tratando los datos con un ordenador. Se ilustran aquí nueve imágenes de una cabeza humana, reconstruidas a partir de una sola colección de datos tridimensionales, que muestran principalmente la distribución de los protones contenidos en el agua y en los lípidos de la sangre y los tejidos. Las tres imágenes de la izquierda muestran secciones transversales a tres niveles: a través de la parte central del cerebro (*arriba*), a nivel del ojo (*centro*) y justo debajo del ojo (*abajo*). Las tres imágenes centrales muestran una sección longitudinal practicada en la línea media (*centro*) y secciones paralelas situadas aproximadamente tres centímetros a izquierda y derecha. Las tres imágenes de la derecha ilustran la cabeza desde la parte frontal, en el punto más ancho del cerebro (*arriba*) y en dos regiones posteriores. Las imágenes, obtenidas en el laboratorio de RMN del Hospital General de Massachusetts, se generaron con un aparato de RMN construido por la Technicare Corporation, de Ohio, que disponía de un imán de 0,15 tesla (1500 gauss).

Estos valores, que se sitúan en la banda de radiofrecuencia del espectro electromagnético, al estar muy por debajo de los de los rayos X o incluso de la luz visible, carecen de la fuerza necesaria para romper las moléculas de los sistemas vivos.

Queda claro, por tanto, que mediante una selección apropiada de la frecuencia pueden "sintonizarse" especies nucleares específicas y observar su respuesta individual. Sin embargo, todas las imágenes médicas de RMN producidas hasta el momento se han obtenido utilizando solamente la resonancia de los núcleos de hidrógeno. Los demás núcleos tienen una menor sensibilidad intrínseca en RMN y se hallan en el material biológico en concentraciones mucho menores.

Desde el punto de vista mecánico-cuántico la desviación del vector de magnetización global de un agregado de núcleos, respecto de su posición de equilibrio, equivale a la transición desde un estado de energía bajo hasta otro más alto. Dicha transición sólo tiene lugar cuando la energía de los cuantos del campo de radiofrecuencia equivale exactamente a la diferencia de energía magnética entre los dos estados energéticos. La posibilidad de generar fácilmente imágenes por RMN en medicina a partir de la resonancia de los núcleos de hidrógeno representa una situación afortunada, ya que el cuerpo humano está formado en un 75 por ciento por agua, y cada molécula de agua tiene dos núcleos de hidrógeno. Además, sabemos que la distribución del agua, y la de varias otras moléculas pequeñas, ricas en hidrógeno (lípidos, por ejemplo), se altera en muchos estados patológicos.

El ángulo de desplazamiento entre el vector de magnetización nuclear *M* y la dirección del campo magnético estático aumenta de forma continua durante todo el tiempo en el que se aplica el campo en rotación a la muestra, y la tasa de incremento depende de la potencia del campo. Se llama pulso de 90 grados al que es suficientemente largo y fuerte para inclinar *M* desde su posición inicial hasta que gire justamente en el plano *x-y*. Pero, ¿cómo sabemos que se ha producido una inclinación? Inmediatamente después de aplicarse un pulso de 90 grados, el vector

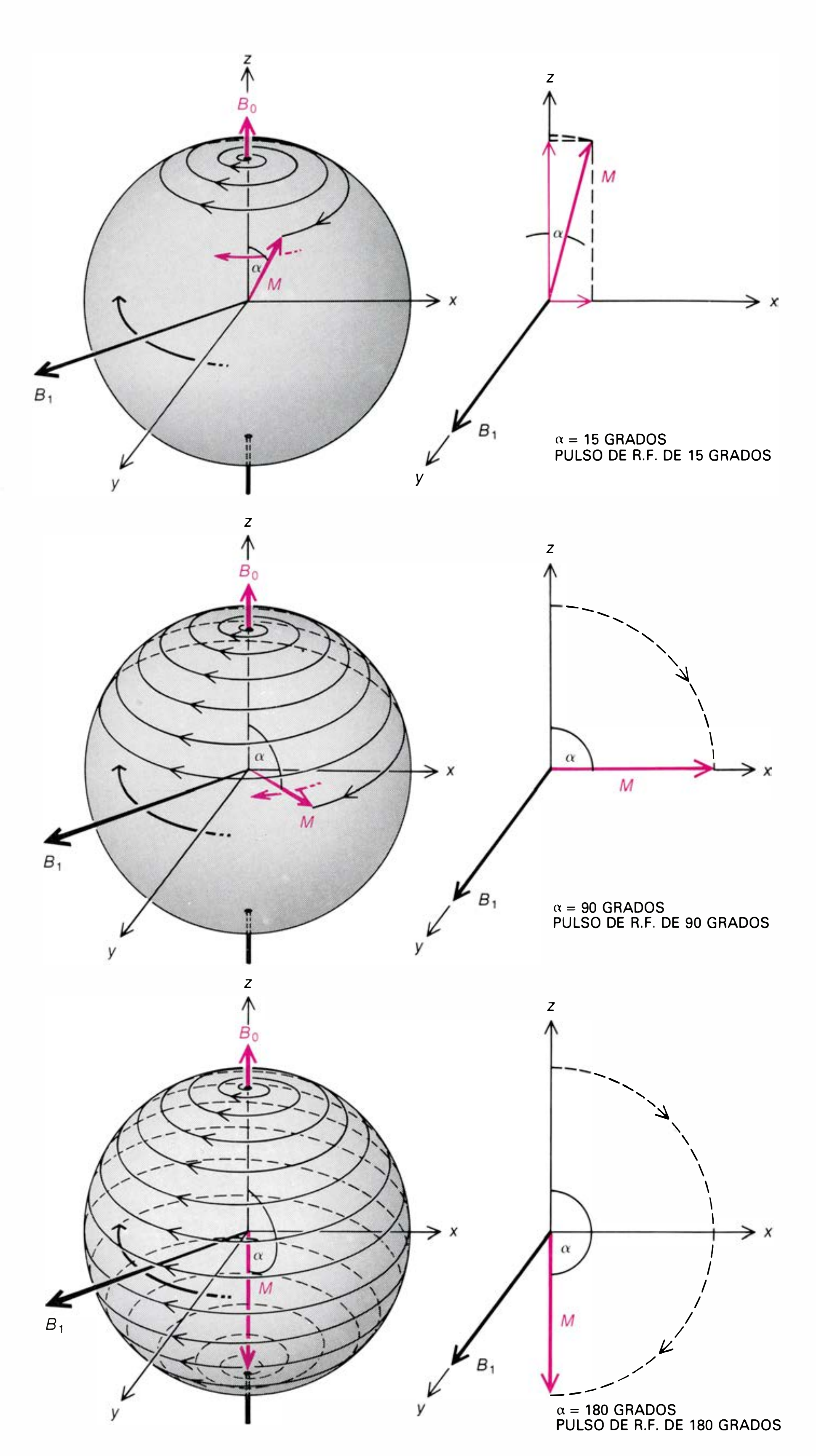


EN UN CAMPO MAGNETICO, B_0, los núcleos con espín no nulo se orientan con las líneas del campo, que apuntan en la dirección *z* (*izquierda*). Se genera así, en el interior de la muestra, un momento magnético neto, *M*, alineado con el campo. En términos mecánico-cuánticos esa alineación comporta la creación de niveles múltiples de energía magnética. En el caso de los núcleos con espín 1/2, donde se incluyen los protones, se establecen dos niveles, alineación en el mismo sentido del campo o en el sentido inverso, con un pequeño exceso de núcleos en el estado energético inferior. Podemos hacer que el vector *M* gire alrededor de la dirección del campo magnético como si se tratase de una peonza. Para ello se aplica (*derecha*) un segundo campo magnético, B_1, con una componente de rotación en el plano *x-y* en ángulo recto con el campo estático (que no gira). *M* sólo se inclina hacia el plano *x-y* si la frecuencia del campo en rotación coincide exactamente con la frecuencia de Larmor (la frecuencia de resonancia natural) de los núcleos. B_1 se produce colocando alrededor de la muestra una bobina que aplica una energía de radiofrecuencia al sistema de espín. En términos mecánico-cuánticos, los cuantos de energía de radio que coinciden con la distancia que separa los niveles de energía provocan el "salto" de algunos núcleos del nivel energético inferior al más alto.

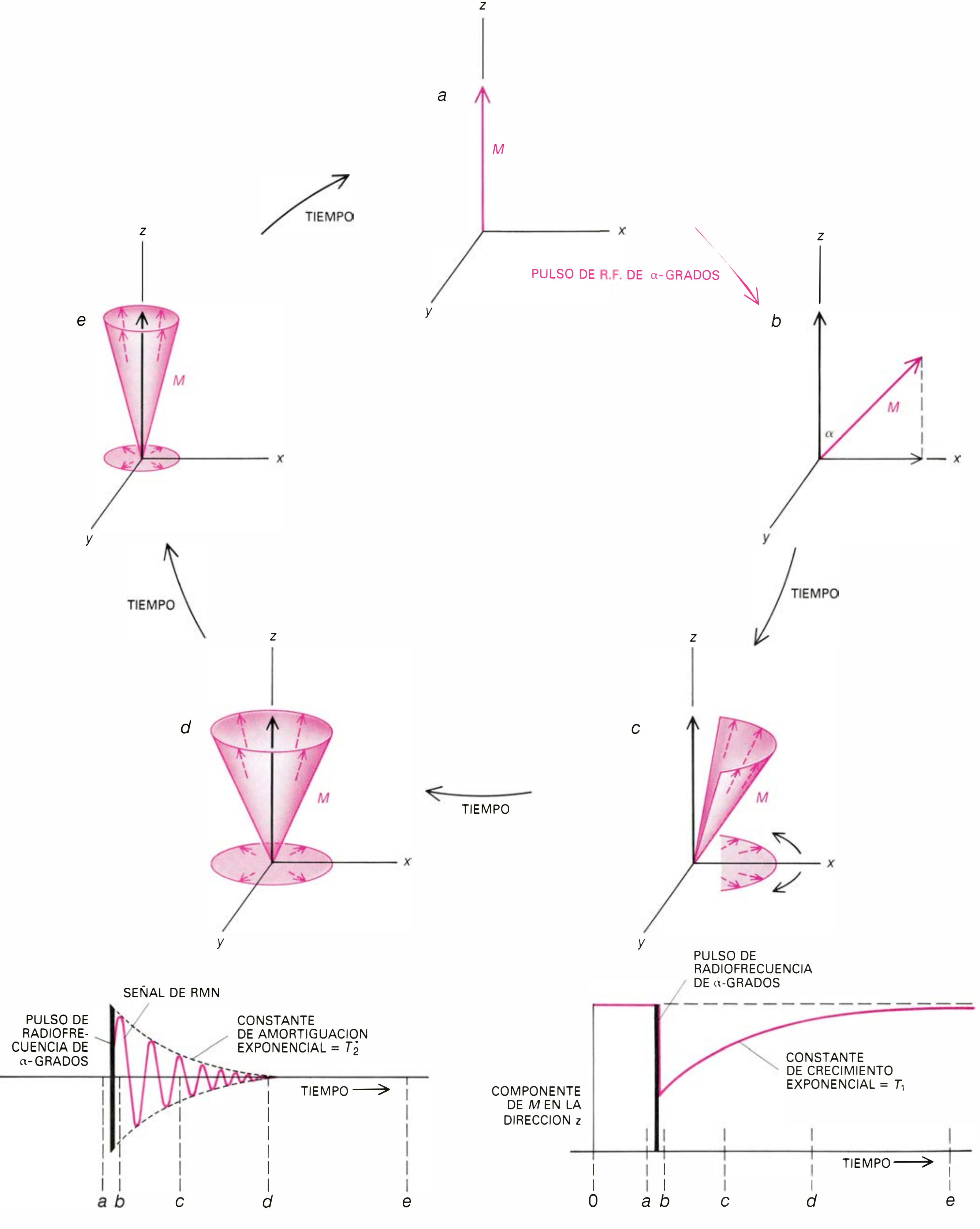
de magnetización continúa girando libremente en el plano x-y, de forma que genera una pequeña fuerza electromotriz, detectable por la misma bobina que transmitió el impulso o por otra bobina receptora. La señal emitida se denomina señal de inducción libre, o amortiguación de inducción libre. En términos mecánico-cuánticos la señal se produce al caer los núcleos del nivel energético excitado al estado fundamental, o nivel de menor energía.

Finalizado el pulso de excitación, el vector de magnetización nuclear acaba por volver a su posición original a lo largo del eje z. El retorno al equilibrio se caracteriza por dos tiempos principales de "relajación": T_1 y T_2. Veamos primero el T_2. Recibe el nombre de tiempo de relajación de espín-espín, o tiempo de relajación transversal. Los fenómenos de relajación de espín-espín afectan a la duración natural de la señal de inducción libre, durante la cual los diversos componentes de magnetización en el plano x-y se mantienen más o menos en fase. Cuando acaba el pulso de excitación, los núcleos no sólo "perciben" el campo estático externo, sino también los campos locales asociados con las propiedades magnéticas de los núcleos vecinos, de forma que van adquiriendo un intervalo de frecuencias precesionales ligeramente diferentes, lo que provoca el desfase de la señal de inducción libre. En un líquido, donde átomos y núcleos están en continuo movimiento, los campos magnéticos internucleares responsables de la relajación de espín-espín tienden a cancelarse; la señal decae mucho más lentamente de lo que sucedería en un sólido, en donde los núcleos se mantienen esencialmente fijos en el espacio. En un líquido puro, T_2 puede alcanzar una magnitud de varios segundos. En un sólido no suele durar más de unos pocos microsegundos; la verdad es que la propia rapidez con que decae la señal impide, por lo general, que se la detecte.

Si el campo magnético estático fuese perfectamente uniforme, nos bastaría medir la velocidad de amortiguación de la señal de inducción libre para determinar el valor de T_2. No obstante, los campos generados por imanes reales casi nunca son perfectos. Hasta las sutilísimas imperfecciones de los mejores imanes utilizados para espectroscopía de RMN hacen que la señal de inducción libre decaiga con mayor rapidez de lo que acontecería en presencia de un campo magnético perfectamente homogéneno. La constante de tiempo que define la velocidad real de amorti-



EL ANGULO DE PRECESION, α, aumenta continuamente mientras se aplica el campo magnético en rotación, B_1. Sin embargo, la frecuencia de precesión se mantiene constante, pues depende de propiedades intrínsecas de los núcleos y de la fuerza del campo estático, B_0. Al pulso de radiofrecuencia necesario para desviar el vector del momento magnético neto M un ángulo de 15 grados se le denomina pulso de 15 grados. Los pulsos de 90 y 180 grados provocan incrementos correspondientes en el ángulo de precesión. Un observador imaginario que girase con el vector de magnetización a la frecuencia de Larmor (*derecha*) vería el aumento del ángulo de precesión como una simple rotación de M alrededor del campo aplicado, B_1, que le parecería estacionario. Siempre que existe un componente neto de M en el plano x-y se genera una fuerza electromotriz detectable en una bobina colocada alrededor de la muestra, que puede ser la misma que emite la señal de radio. Esta fuerza electromotriz es el origen de la señal de RMN.



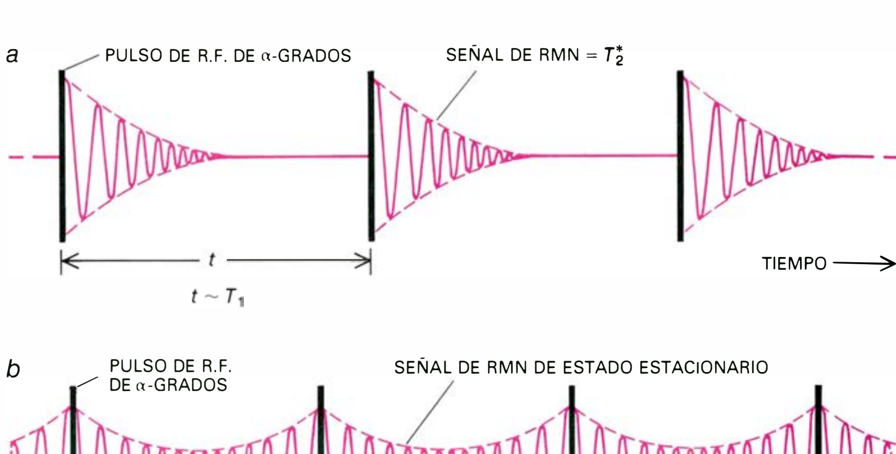
EN UN EXPERIMENTO DE RMN POR IMPULSOS se observa la señal emitida después de desconectar la energía de radiofrecuencia que induce la precesión. En estos diagramas se supone que el marco de referencia gira a la frecuencia media de Larmor. Cuando un pulso de radiofrecuencia ha inclinado el vector del momento magnético neto, M, un cierto ángulo, α, existirá una componente de M en el plano x-y (*b*). Por un breve instante, la señal de RMN alcanza su valor máximo (*curva de la parte inferior izquierda*). Los núcleos empiezan a girar a velocidades diferentes como consecuencia de las ligeras desigualdades del campo magnético y de las interacciones magnéticas que se producen entre ellos mismos; la componente neta de M en el plano x-y disminuye, y la amplitud de la señal cae exponencialmente según la constante de tiempo T^*_2 (*c, d*). Si el campo magnético B_0 es perfectamente uniforme, el tiempo de amortiguación de la señal será más largo y la constante de tiempo será T_2, denominada tiempo de relajación de espín-espín. Simultáneamente, la componente longitudinal de magnetización aumenta a medida que M vuelve a su posición de equilibrio, alineado con el eje z (*e*). Esta relajación, designada T_1 (tiempo de relajación de espín-red), mide el tiempo necesario para que el sistema de espín vuelva a su equilibrio térmico (*curva de abajo, derecha*).

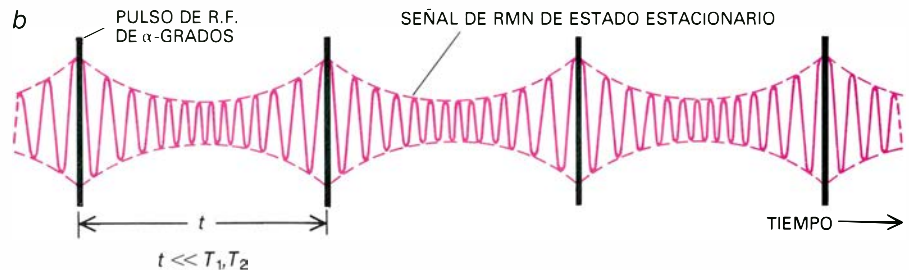
guación de la señal en un campo imperfecto se designa por T^*_2 ("T dos estrella") para distinguirla del verdadero tiempo de relajación T_2.

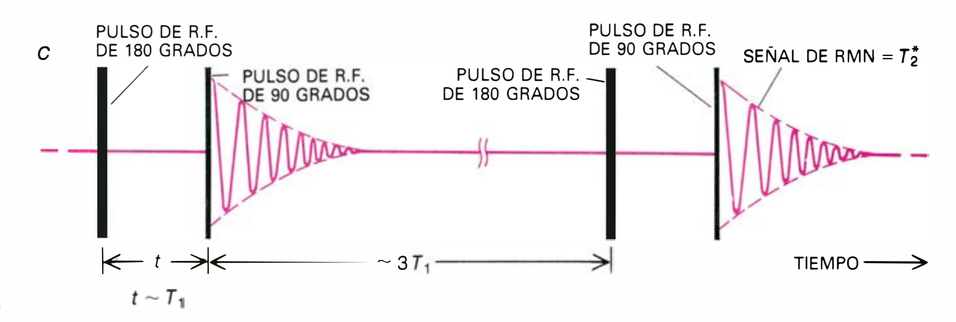
A pesar de todo ello, podemos determinar el valor intrínseco de la T_2 para una muestra aunque esté sometida a un campo imperfecto, ya que las desigualdades del campo magnético son constantes, lo que permite identificarlas y cancelarlas. Se recoge la señal en forma de un "eco de espín" o de una serie de ecos, aplicando un perfil especial de pulso de radiofrecuencia, conocido como la secuencia de pulsos Carr-Purcell. En dicha secuencia, la señal inicial de amortiguación de inducción libre y cada uno de los ecos de espín individual decaen con una constante de tiempo T^*_2, aunque las alturas del pico de los sucesivos ecos del espín decaen con una constante de tiempo igual al valor intrínseco de T_2 en la muestra.

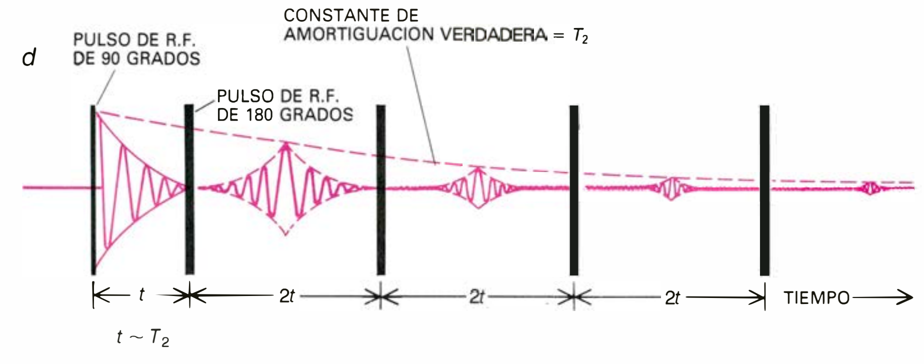
El otro tiempo de relajación, T_1, se conoce como el tiempo de relajación del espín-red, o tiempo de relajación longitudinal. T_1 es característico del tiempo necesario para que el sistema de espín recupere el equilibrio térmico con su entorno (la "red") una vez concluido el pulso de excitación. En términos mecánico-cuánticos, los campos magnéticos fluctuantes de los núcleos que forman la red cristalina deben tener los componentes de frecuencia apropiados para estimular transiciones desde el nivel de energía magnética superior hasta el estado fundamental. En los sólidos o muestras de material a bajas temperaturas, en donde los átomos y moléculas se desplazan muy poco, no existirán muchos componentes con la frecuencia correcta, por lo que T_1 puede durar horas. Para los protones de líquidos puros, simples, tales como el agua destilada, T_1 y T_2 son aproximadamente iguales (unos pocos segundos), lo que constituye una indicación de la dependencia mutua de los tiempos de relajación sobre los campos magnéticos internos fluctuantes.

En consecuencia, en materiales líquidos y licuiformes, la relación de T_2 a T_1 tiende a aproximarse a la unidad. En los sólidos esa proporción es muy pequeña. Las interacciones magnéticas internucleares ("dipolo-dipolo") que hacen que T_2 sea extremadamente reducida en los sólidos pueden cancelarse recurriendo a ciclos complejos de pulsos múltiples. Sin embargo, dichos ciclos no se han empleado todavía para la formación de imágenes por RMN en medicina; sólo las regiones licuiformes de los materiales biológicos dan señales apreciables.









DIVERSOS PERFILES DE IMPULSOS, que realzan los diferentes aspectos de la señal de RMN inducida y que se aprovechan para variar el contraste de las imágenes de RMN. En una secuencia de pulsos de "saturación-recuperación" (*a*) se aplica una serie de pulsos de 90 grados con una separación entre cada pulso, *t*, de mayor duración que el tiempo de amortiguación T_2 y aproximadamente igual a T_1. Si se modifica el valor de *t*, las variaciones de T_1 en las diferentes partes de la muestra aparecerán como diferencias de intensidad de imagen, pudiendo generarse de esta forma un "mapa de T_1". Si *t* es mucho más breve que T_2, la señal no decaerá hasta cero entre los sucesivos pulsos, y ello creará una condición conocida como precesión libre de estado estacionario, o PLEE (*b*). En este caso el contraste de la imagen puede modificarse variando el ángulo de pulso, α. En la secuencia de pulsos de "inversión-recuperación" (*c*) se invierte el vector de magnetización aplicando primero un pulso de 180 grados y, a continuación, un pulso "lector" de 90 grados. Esta secuencia también acentúa las variaciones de T_1. La secuencia de "eco de espín Carr-Purcell" (*d*) consiste en un pulso de 90 grados seguido por una serie de pulsos de 180 grados. En este caso las imágenes resultantes dependen mucho del tiempo de relajación de espín-espín, T_2.

Hemos visto ya cómo determinar el tiempo de relajación T_2 intrínseco del material. También puede determinarse el valor intrínseco de T_1. Se consigue aplicando un segundo pulso después de la amortiguación de la señal, pero antes de que se haya alcanzado totalmente la relajación de espín-red. El segundo pulso inducirá otra señal de inducción libre, aunque en este caso la amplitud de la misma será menor. La razón es que la potencia de la señal inducida es proporcional al valor que la componente longitudinal de la magnetización a lo largo de la dirección z del campo presenta justo antes de la aplicación del

pulso. Antes del pulso inicial todos los núcleos están alineados con el campo, y, por tanto, la señal inducida es máxima. Sin embargo, el segundo pulso se aplica cuando la componente de magnetización está todavía de regreso a su posición alineada, de suerte que la señal inducida no es tan fuerte como la original. Puesto que la componente de magnetización vuelve a su valor de equilibrio de forma exponencial según una constante de tiempo T_1, la amplitud respectiva de las dos señales representa una medida de dicha constante.

Las variaciones en el valor de T_1 pueden aprovecharse en la formación de imágenes por RMN para incrementar el contraste entre diferentes regiones de muestras de tejido blando. Las diferencias en los valores de T_1 se realzan mediante secuencias especiales de pulsos, por ejemplo la secuencia de "saturación-recuperación". En este caso, se expone la muestra a un tren de pulsos de radiofrecuencia con un espaciado constante entre pulsos. Suponiendo que en cada intervalo entre pulsos se produzca una relajación espín-espín total (T_2), la amplitud de las señales de amortiguación de inducción libre alcanzará un estado estacionario dependiente tanto del valor de T_1 como de la densidad de los núcleos licuiformes que contribuyen a la señal. La relación con la fuerza de la señal es tal que, para las zonas de la imagen donde T_1 es largo en comparación con el intervalo entre pulsos, la intensidad de la imagen será débil.

Se dice que este procedimiento proporciona imágenes ponderadas en T_1. Si se obtienen dos o más imágenes con un diferente espaciado entre los pulsos, puede calcularse a partir de los datos un valor T_1 correspondiente a cada elemento de imagen (cada pixel), independiente de la densidad de espín. El resultado es un "mapa de T_1". Las diferencias de valores de T_1 así identificados son extremadamente útiles en la formación de imágenes de RMN de protón, puesto que las diferencias de T_1 de protón entre diversos tejidos blandos son muy superiores a las correspondientes diferencias en la densidad de protón móvil.

Si los pulsos se aplican a una velocidad tal que la señal no decae a cero entre pulsos sucesivos, se produce un nuevo tipo de estado estacionario, en el que la intensidad de la señal depende tanto de T_2 como de T_1. Esta secuencia de pulsos recibe el nombre de precesión libre de estado estacionario, o PLEE. Aunque se han podido generar imágenes de excelente calidad utilizando la secuencia de PLEE, no es fácil separar las respectivas contribuciones de T_1 y T_2 a la intensidad de la imagen resultante.

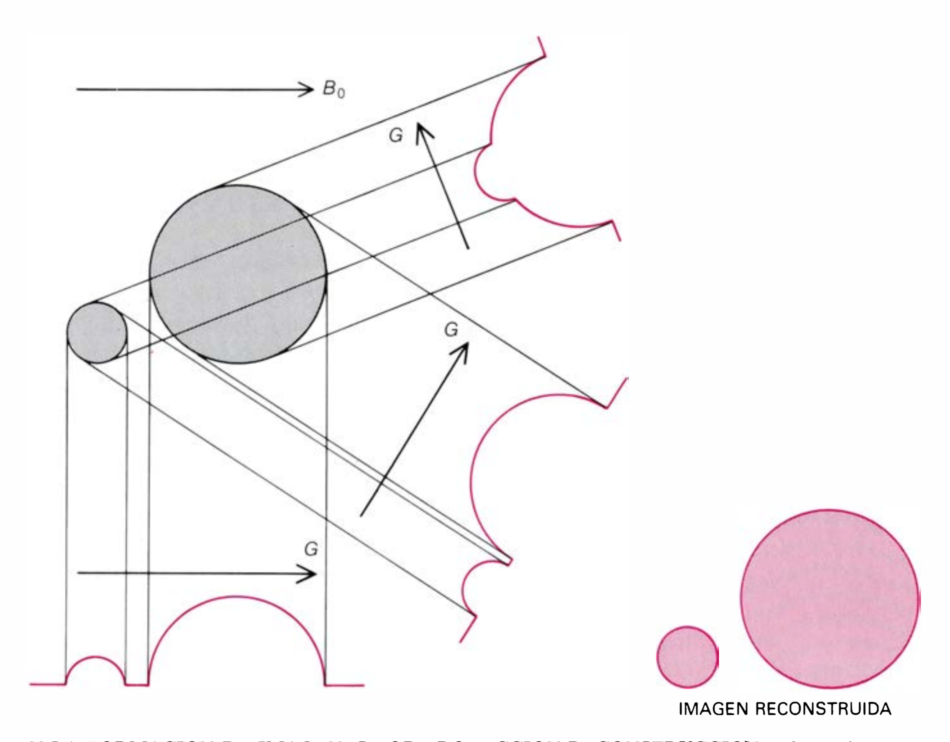
Otro perfil de pulsos adoptado frecuentemente es el de la secuencia de "inversión-recuperación", que se parece a la secuencia de saturación-recuperación en el sentido de que se realzan así las variaciones de T_1 de la muestra. En este perfil se invierte primero el vector de magnetización global al aplicar un pulso de 180 grados a la muestra. La relajación T_1 procede durante un intervalo seleccionado de pulsos, después del cual se aplica un pulso "lector" de 90 grados. La señal de inducción libre resultante del pulso lector sirve para generar la imagen.

La secuencia de inversión-recuperación proporciona una imagen de mayor contraste que las secuencias de saturación-recuperación, pero a cambio de un mayor tiempo de formación de imágenes o de una reducción de la resolución espacial. El motivo es que, para evitar errores en las determinaciones de T_1, debe dejarse transcurrir un tiempo de demora equivalente, por lo menos, a tres veces el valor de T_1 antes de proceder a la repetición del par de pulsos de 180 y 90 grados. A partir de los perfiles de pulsos de inversión-recuperación también pueden obtenerse mapas de T_1. La secuencia de pulsos de eco de espín antes citada sirve para crear imágenes que dependen principal o únicamente del tiempo de relajación de espín-espín, T_2.

Por tanto, seleccionando una secuencia apropiada de pulsos puede conseguirse que la imagen de RMN refleje uno o más de los diversos parámetros de RMN inherentes al tejido sometido a examen. Dichos parámetros, a su vez, son sensibles al entorno fisicoquímico de los núcleos y subrayan la convicción de que la formación de imágenes de RMN puede desempeñar un gran papel en la detección precoz de enfermedades y en el control de su curso.

LOS METODOS PARA CREAR UNA IMGEN POR RMN requieren una "codificación" espacial de la señal de RMN. La ilustración presenta una muestra de agua en dos cilindros, vista desde arriba. La señal de RMN de hidrógeno obtenida en espectroscopía por RMN (*izquierda*) aparece como una espiga única después de la transformación de Fourier, proceso matemático que transforma una curva que representa la fuerza de la señal en función del tiempo en otra que representa la fuerza de la señal en función de la frecuencia. Si a continuación se le añade al campo estático original, B_0, un gradiente de campo magnético lineal, G, la señal inducida después de su transformación de Fourier adopta la forma de una curva que representa el contorno de la muestra (*derecha*). El área comprendida bajo las curvas representa el número total de protones en la muestra, y por tanto es la misma en ambas. Ilustración de Walken Graphics.

Los métodos descritos para la manipulación de la respuesta de RMN no sólo se aplican en la formación de imágenes de RMN, sino también en espectroscopía de RMN, donde la señal inducida representa la suma de la respuesta de RMN procedente de la totalidad de la muestra de material. No obstante, si la señal nuclear procede de toda la muestra, ¿cómo puede codificarse con información espacial? Ya en 1951 Robert Gabillard, de la École Normale Supérieure de París, se planteó la cuestión inversa: observó que la señal de RMN podría distorsionarse de acuerdo con la forma y el tamaño de la muestra. El fenómeno se atribuyó, correctamente, a las desigualdades del campo magnético estático. El grado de distorsión depende de la porción de muestra que esté sometida a las zonas no uniformes del campo y de la magnitud de las desigualdades.



EN LA FORMACION DE IMAGENES POR PROYECCION-RECONSTRUCCION se hace girar un gradiente de campo magnético para obtener "instantáneas" de la muestra a diferentes ángulos, cubriendo un arco de por lo menos 180 grados. Partiendo de ese grupo de datos un ordenador puede reconstruir una imagen transversal de la muestra. El método es análogo al de la tomografía computarizada de rayos X.

De ahí el sostenido esfuerzo que se está realizando en el campo de la instrumentación de RMN por eliminar la influencia de la forma de la muestra sobre la señal de RMN, fabricando para ello imanes con campos cada vez más homogéneos y estables. Este problema puede haber contribuido a retrasar la implantación del método de formación de imágenes por RMN, puesto que en esta técnica debe crearse un campo no uniforme, aunque de forma controlada, generalmente por superposición sobre el mismo de un gradiente de campo magnético lineal.

El motivo principal que ha impulsado el perfeccionamiento del campo estático de los imanes es el empeño puesto por los espectroscopistas de RMN en medir el sutil "desplazamiento químico" de las muestras que contienen moléculas de estructura compleja. Uno esperaría que todos los núcleos de una misma especie, sometidos a un campo magnético homogéneo, ofreciesen la misma frecuencia de resonancia; en otras palabras, que apareciera un pico único y estrecho en el espectro de frecuencias de RMN. Ello es cierto para los núcleos contenidos en muestras de moléculas muy simples (como por ejemplo los núcleos de hidrógeno en el agua pura). Sin embargo, para moléculas más complejas, el campo magnético que rodea a algunos núcleos queda sutilmente alterado por la acción de "corrientes de pantalla" asociadas a la distribución de los electrones alrededor de los átomos adyacentes. Estas alteraciones provocan ciertos desplazamientos de la frecuencia de resonancia que son característicos de determinadas conformaciones moleculares, y que facilitan la determinación directa de la estructura química.

Estos desplazamientos químicos son muy pequeños, midiéndose en partes por millón con respecto a la fuerza del campo magnético estático; ello explica por qué existe tanto interés en la construcción de imanes con un campo altamente uniforme. En la actual generación de sistemas de imagen por RMN, la insuficiente uniformidad del campo magnético estático no permite observar desplazamientos químicos con un grado de exactitud útil. En cualquier caso, el propio procedimiento de formación de imágenes (específicamente la aplicación de los gradientes de campo) impide muchas veces la extracción de ese tipo de información.

A Paul C. Lauterbur, de la Universidad estatal de Nueva York en Stony Brook, debemos la publicación de la primera imagen obtenida por RMN. En un artículo aparecido en 1973 presentó imágenes de dos tubos capilares llenos de agua, obtenidas con un espectrómetro de RMN modificado. La muestra podía considerarse un medio que acoplaba el campo de radiofrecuencia con el campo del gradiente magnético, por lo que Lauterbur denominó a su método de formación de imágenes zeugmatografía, del griego *zeugma*, vocablo indicativo de aquello que une, como el yugo.

Los sistemas de formación de imágenes por RMN pueden diseñarse de forma que reciban datos de un único punto, de una línea, de un plano o de la totalidad del volumen tridimensional de la muestra. En los métodos de exploración o barrido por puntos o líneas se generan las imágenes por desplazamiento electrónico del punto o línea seleccionados a través de la muestra en un rastreo, o perfil secuencial. Pero esos métodos han sido casi totalmente reemplazados por otros bidimensionales y tridimensionales, mucho más eficaces.

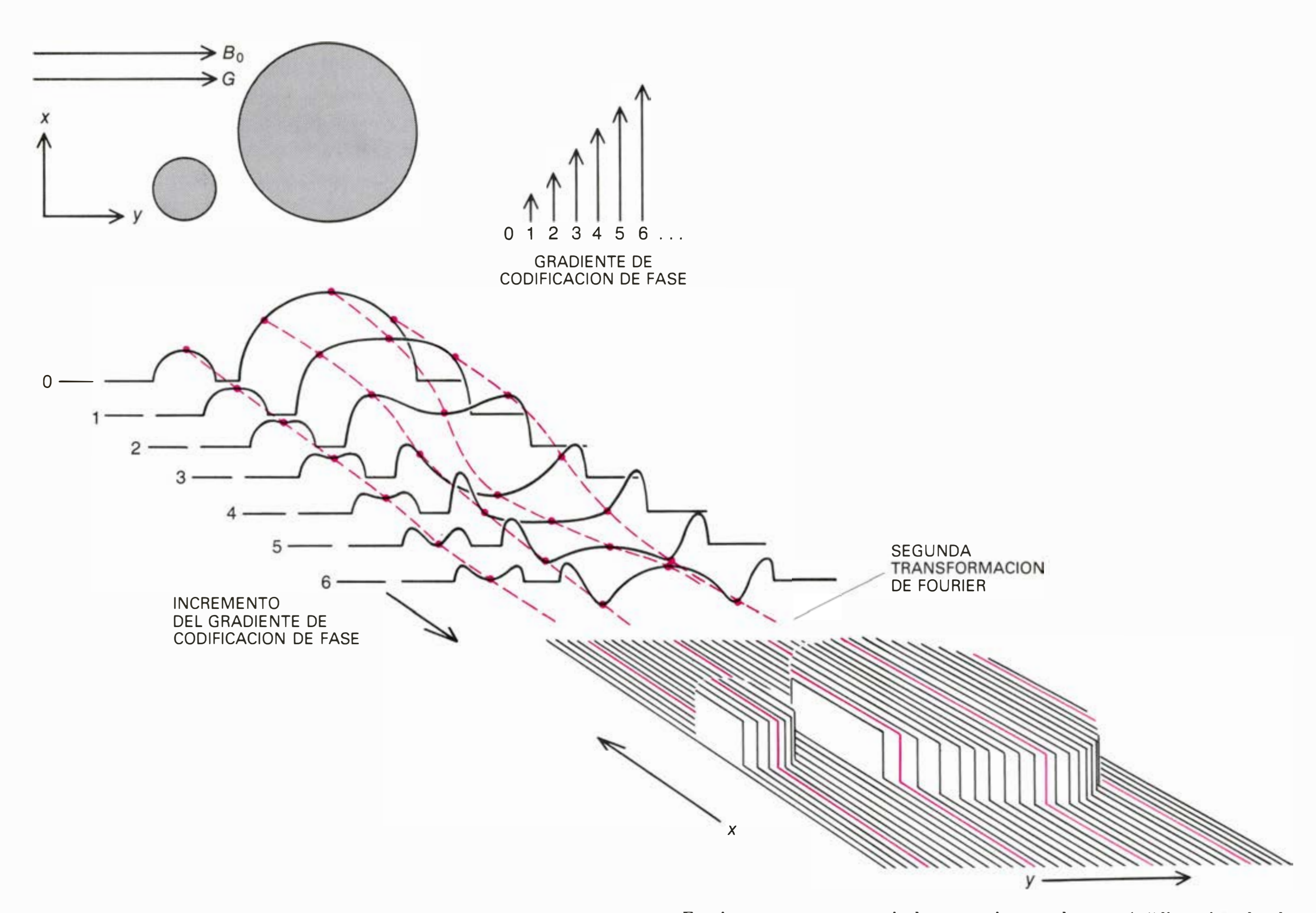
La mayoría de los métodos bi- y tridimensionales requieren la aplicación de un gradiente de campo magnético lineal para generar una dimensión de información espacial. La resolución en la segunda y tercera dimensiones se logra aumentando secuencialmente las magnitudes o direcciones del gradiente a través de un juego de valores predeterminados. Hay, empero, otras posibilidades. En este sentido, la zeugmatografía de marco en rotación, descrita por primera vez por David Hoult, del National Institute of Dental Research, es un método único, en cuanto utiliza un gradiente en el campo de radiofrecuencia y otro en el campo magnético estático. Peter Mansfield, de la Universidad de Nottingham, ha desarrollado

un método de muy alta velocidad, la formación de imágenes ecoplanares, que genera una imagen completa y resuelta espacialmente después de un único pulso de radiofrecuencia.

Las primeras imágenes producidas por Lauterbur utilizaron los algoritmos de ordenador para reconstrucción de la imagen que se usan en el barrido de TC. Si se coloca una muestra de agua en un campo magnético homogéneo, el espectro de frecuencia de RMN de los núcleos de hidrógeno en las moléculas del agua es una línea única y estrecha. Si el campo magnético es perfectamente uniforme, la forma de la línea no depende de la geometría de la muestra. Si se superpone entonces un gradiente de campo magnético lineal, los núcleos resonantes en un lado de la muestra percibirán un campo magnético total más débil que los situados en el otro lado, lo cual producirá una distribución lineal de frecuencias de Larmor a través de la muestra. Se podrá luego someter la señal de amortiguación de inducción libre a una transformación de Fourier, procedimiento matemático que convierte los datos de una curva representativa de la potencia de la señal en función del tiempo en otra gráfica que representa la potencia de la señal en función de la frecuencia.

Resulta un espectro ensanchado hasta una forma correspondiente a la proyección unidimensional de la potencia de la señal de RMN sobre el eje de frecuencia. Haciendo girar electrónicamente el gradiente de campo magnético se obtiene una proyección desde un ángulo ligeramente diferente. El análisis por ordenador de muchas de estas proyecciones reconstruye la geometría de la muestra. En la aplicación bidimensional de la técnica, la dirección del gradiente se hace girar en un plano. En la extensión tridimensional del método, el gradiente gira en el espacio tridimensional, al menos a través de media esfera.

La transformación de Fourier para imágenes bidimensionales, o zeugmatografía de Fourier, fue demostrada, en 1975, por A. Kumar, D. Welti y R. R. Ernst, del Instituto Politécnico Federal Suizo. Su método debe más a la espectroscopía por RMN que a los algoritmos de reconstrucción de TC, puesto que en este caso, para codificar espacialmente la señal, se obtiene tanto información de amplitud como de fase. El primer paso requiere de nuevo la generación de una proyección unidimensional, aunque ahora se aplica un gradiente de codificación de fase justo antes de activar el gradiente original. El gradiente de codificación de fase se aplica en ángulo recto al gradiente original, aumentando sucesivamente su duración (o amplitud) desde el valor cero, en lugar de girar progresivamente el gradiente, como se hacía en la forma original de la zeugmatografía. Las proyecciones codificadas en fase se almacenan en orden creciente de magnitud del gradiente de codificación de fase, transformándose por métodos de Fourier los correspondientes puntos



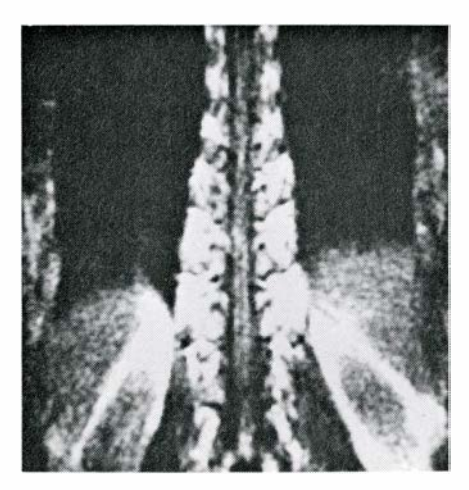
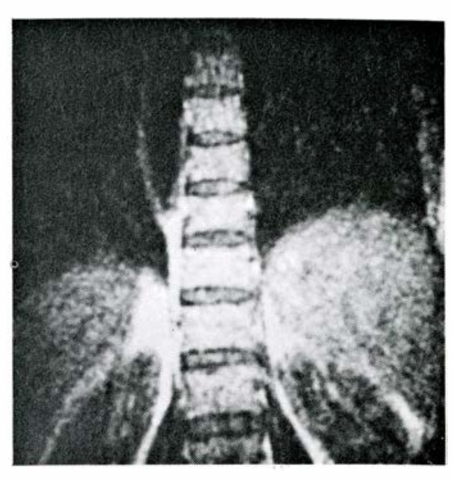
EN LA ZEUGMATOGRAFIA DE FOURIER se aplica otro gradiente de campo magnético durante un breve período de tiempo justo antes del gradiente usual, *G*. Este nuevo gradiente, conocido por gradiente de codificación de fase, se aplica en ángulo recto respecto al original, aumentándose su magnitud, desde cero, en una serie de pequeños incrementos. El efecto de los dos gradientes sobre la señal de RMN es tal que después de la transformación de Fourier se generan una serie de proyecciones cada vez más "distorsionadas de fase". A continuación, para generar la imagen final, los correspondientes puntos de cada una de las proyecciones (*líneas discontinuas de color*) se someten a una segunda transformación de Fourier. Zeugmatografía procede del griego, *zeugma*, vocablo indicativo de lo que une (el yugo). Se refiere a cómo la muestra acopla los campos de radiofrecuencia y de gradiente magnético.

procedentes de cada proyección una segunda vez para generar la imagen final [*véase la ilustración de la página anterior*].

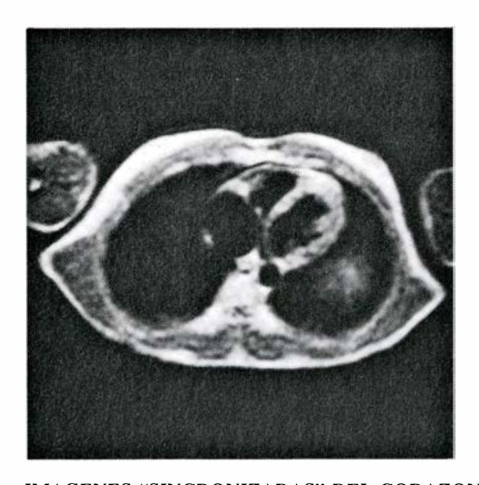
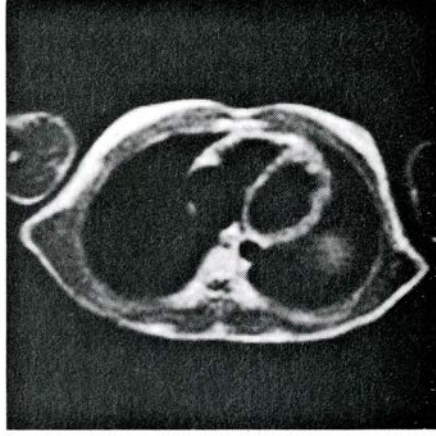
Por su propia naturaleza, la resonancia magnética nuclear es un fenómeno tridimensional. Las señales de RMN proceden normalmente del volumen total de material incluido en el interior de las bobinas de transmisión y recepción; se precisa, pues, un considerable grado de ingenio para reducir a puntos, líneas o planos definidos la región de la que proceden las señales. En un método, denominado de irradiación selectiva, se aplica al efecto un pulso de radiofrecuencia especialmente diseñado, de suerte que sólo comprenda una banda muy estrecha de frecuencias. Unicamente los núcleos situados dentro de una sección perpendicular a la dirección de un gradiente de selección de plano exhibirán frecuencias resonantes correspondientes a las del pulso de radiofrecuencia. En consecuencia, sólo se irradia una rebanada aislada y delgada de material. El espesor o posición del plano puede alterarse variando electrónicamente la anchura o desplazamiento de la banda de frecuencias del espectro de irradiación.

Un segundo método, desarrollado por Waldo S. Hinshaw cuando se hallaba en la Universidad de Nottingham, aplica un campo magnético oscilante para seleccionar un plano determinado. En este método se invierte periódicamente la dirección del gradiente de campo de selección de plano, muchas veces de forma sinusoidal. Existe entonces un único plano, en el fulcro de oscilación, en cuyo interior el campo magnético se mantiene independiente del tiempo; las señales procedentes del exterior de este plano varían de tal modo que no contribuyen a la intensidad de la imagen. Los aquí reseñados son dos de entre los varios métodos posibles para la selección de planos.

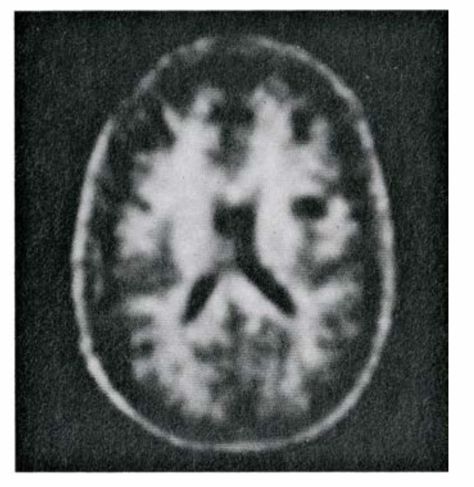
Aunque los métodos de volumen total ofrecen ventajas importantes, también presentan algunos inconvenientes técnicos. La enorme cantidad de información adquirida exige la utilización de un ordenador de gran capacidad de manipulación y acumulación de datos. Por ejemplo, la generación de una imagen tridimensional con 256 puntos en cada dimensión y 256 niveles de intensidad (igual a 8 bits) por cada punto exige un sistema con más de 134 millones de bits de memoria ($256^3 \times 8$). Además, para definir con nitidez los distintos puntos de una matriz tridimensional de datos

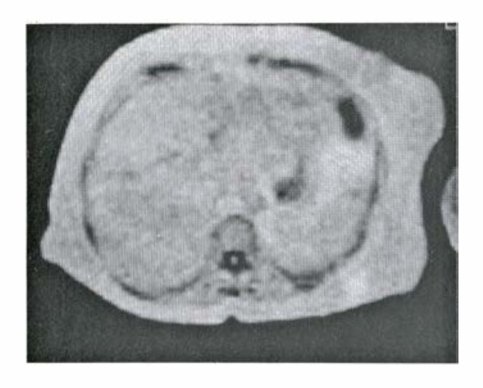
REGION LUMBAR DE LA COLUMNA VERTEBRAL humana a partir de dos reconstrucciones de grupos de datos tridimensionales de RMN. La imagen izquierda muestra la médula espinal en el interior del canal vertebral. En la imagen de la derecha el plano de la misma se ha desplazado unos pocos centímetros para poner de manifiesto los discos intervertebrales. Con un barrido de tomografía computarizada, o mediante cualquier otro método de formación de imagen, resultaría difícil obtener vistas claras de la columna vertebral en una área de esas dimensiones. Las imágenes fueron facilitadas por Technicare.

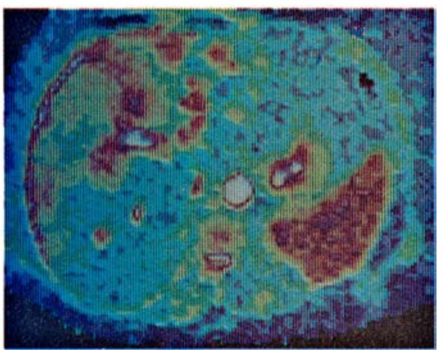



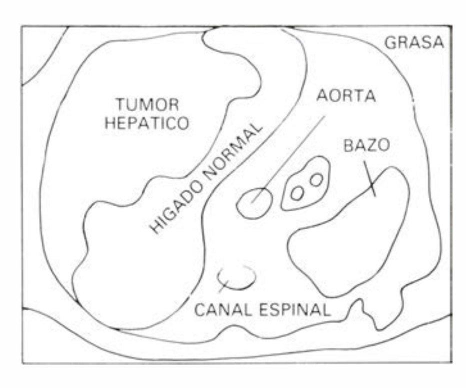
IMAGENES "SINCRONIZADAS" DEL CORAZON, plasmadas, por resonancia magnética nuclear, en secciones transversales del tórax humano. (Las dos formas más pequeñas que aparecen a los lados del tórax son los brazos.) A la izquierda se muestra el corazón al final del sístole, cuando se vacían las cámaras. A la derecha, el final del diástole, cuando se llenan las cámaras. Las imágenes, facilitadas por Technicare, se obtuvieron por sincronización del registro de los datos con la fase del ciclo cardíaco.




IMAGENES DE RMN Y DE TC de rayos X de dos planos transversales del cerebro de un paciente afecto de un tumor cerebral. En la de RMN (*izquierda*), reconstruida a partir de datos tridimensionales, el tumor aparece como una zona circular oscura en el lado derecho del cerebro. En la exploración tomográfica (*derecha*) el tumor no se aprecia en absoluto. Las imágenes, realizadas en este caso por el Hospital General de Massachusetts, se reproducen con permiso del *Journal of Computerized Axial Tomography*.







SECCION TRANSVERSAL DE UN PACIENTE DE CANCER, según dos imágenes de RMN realizadas por James M. S. Hutchison, Francis Smith y John R. Mallard y colaboradores, de la Facultad de Medicina de la Universidad de Aberdeen, utilizando un dispositivo de 0,04 tesla. La imagen en blanco y negro es un simple registro de densidad de protones. La imagen en color es una síntesis de ordenador de una serie de imágenes, obtenidas alterando la secuencia de pulsos para discriminar entre distintos tejidos de acuerdo con las diferencias del tiempo de relajación T_1. Se introdujo luego la codificación de color para realzar la discriminación. El paciente tenía un cáncer de recto que se había extendido al hígado. El tejido hepático normal se representa en azul claro. Las regiones cancerosas tienen un color amarillo-marrón. La médula y canal espinal se aprecian mejor en la imagen de densidad de protones.

como esta se necesitan muchos valores secuenciales de la magnitud o dirección del gradiente de campo. Ello alarga el tiempo de recogida de datos, en particular si se desean obtener mapas de T_1. Por tales razones, es preferible a veces generar un pequeño número de imágenes bidimensionales seleccionadas. Aunque es cierto que normalmente puede obtenerse una imagen bidimensional en menos tiempo que una verdadera imagen tridimensional, esta última puede "diseccionarse" a voluntad en un gran número de cortes, de modo que el tiempo de obtención de una imagen por plano se reduce considerablemente.

La resolución espacial de un grupo de datos tridimensionales suele ser isótropa, o equivalente en las tres dimensiones. Lo que nos permite conseguir cortes bidimensionales de espesor seleccionado, en cualquier posición u orientación, cuando se dispone de datos primarios. Por otro lado, con los datos tridimensionales en la mano pueden detectarse matemáticamente las superficies, lo que faculta al clínico para determinar el volumen de los órganos o de lesiones patológicas.

En la práctica médica hay que considerar muchos factores al seleccionar un determinado método de formación de imágenes, especialmente la escala de tiempo de los movimientos involuntarios del tejido sometido a estudio. Por ejemplo, la cabeza resulta particularmente apropiada para la formación de imágenes tridimensionales, ya que puede mantenerse inmóvil cuanto dure el barrido. En cambio, el corazón, que late incesantemente, requiere un método de formación de imágenes de alta velocidad o bien un método capaz de sincronizar los datos recogidos a lo largo de una serie de ciclos cardíacos.

El aparato necesario para la formación de imágenes por RMN explota la misma tecnología básica desarrollada para la espectroscopía por RMN. Muchos de los primeros experimentos de formación de imágenes se realizaron con espectrómetros de RMN modificados. Podemos mejorar la relación señal/ruido de una imagen de RMN si aumentamos la fuerza del campo magnético estático del aparato. A medida que se incrementa la fuerza del campo, crece también linealmente la frecuencia de precesión de Larmor de los núcleos de la muestra de material a examen, lo que obliga a emplear una radiofrecuencia mayor. El problema que nos plantea este caso es que, a medida que aumenta la frecuencia, tanto las señales transmitidas como las emitidas se absorben con mayor intensidad. Si se pretende tomar imágenes de todo el cuerpo humano, esta atenuación puede convertirse en un factor limitante cuando la frecuencia supera aproximadamente los 15 megahertz, que en las imágenes de protones corresponde a una fuerza de campo de 0,35 tesla (3500 gauss). Aunque en espectroscopía por RMN semejante fuerza se consideraría baja, en el caso de formación de imágenes por RMN el gran volumen de trabajo en el que se precisa uniformidad de campo magnético ha exigido el planteamiento de nuevos enfoques para el diseño de imanes.

Los dos diseños de imán que privan hoy son el imán de temperatura ambiente con núcleo de aire y cuatro bobinas y el imán superconductor de gran calibre enfriado por helio. Los sistemas de RMN con imanes convencionales no superconductores cuestan menos y resultan enteramente satisfactorios para la formación de imágenes de protones de todo el cuerpo a potencias de campo inferiores, aproximadamente, a los 0,2 tesla, para los que el consumo de potencia ronda los 50 kilowatt y las necesidades de refrigeración no resultan prohibitivas. Los imanes superconductores, aunque exigen un mayor desembolso inicial, son de menor coste operativo y generan campos magnéticos más fuertes y más estables que los que se obtienen con imanes no superconductores. Por tanto, son preferibles los imanes superconductores si se pretende tomar imágenes con núcleos que no sean el hidrógeno, lo que implica utilizar campos significativamente más potentes. Por ejemplo, la observación del fósforo 31 a 15 megahertz exige una potencia de campo de 0,87 tesla. Un componente crítico de un aparato de formación de imágenes de RMN que normalmente no se encuentra en el aparato de espectroscopía de RMN es el sistema de bobina para gradiente, necesario para crear gradientes de campo lineal cuya magnitud a veces tiene que alterarse muy rápidamente. La colisión entre las contrapuestas exigencias de gradientes de gran magnitud, por un lado, y la rapidez de variación, por otro, han estimulado la creación de muchos diseños originales.

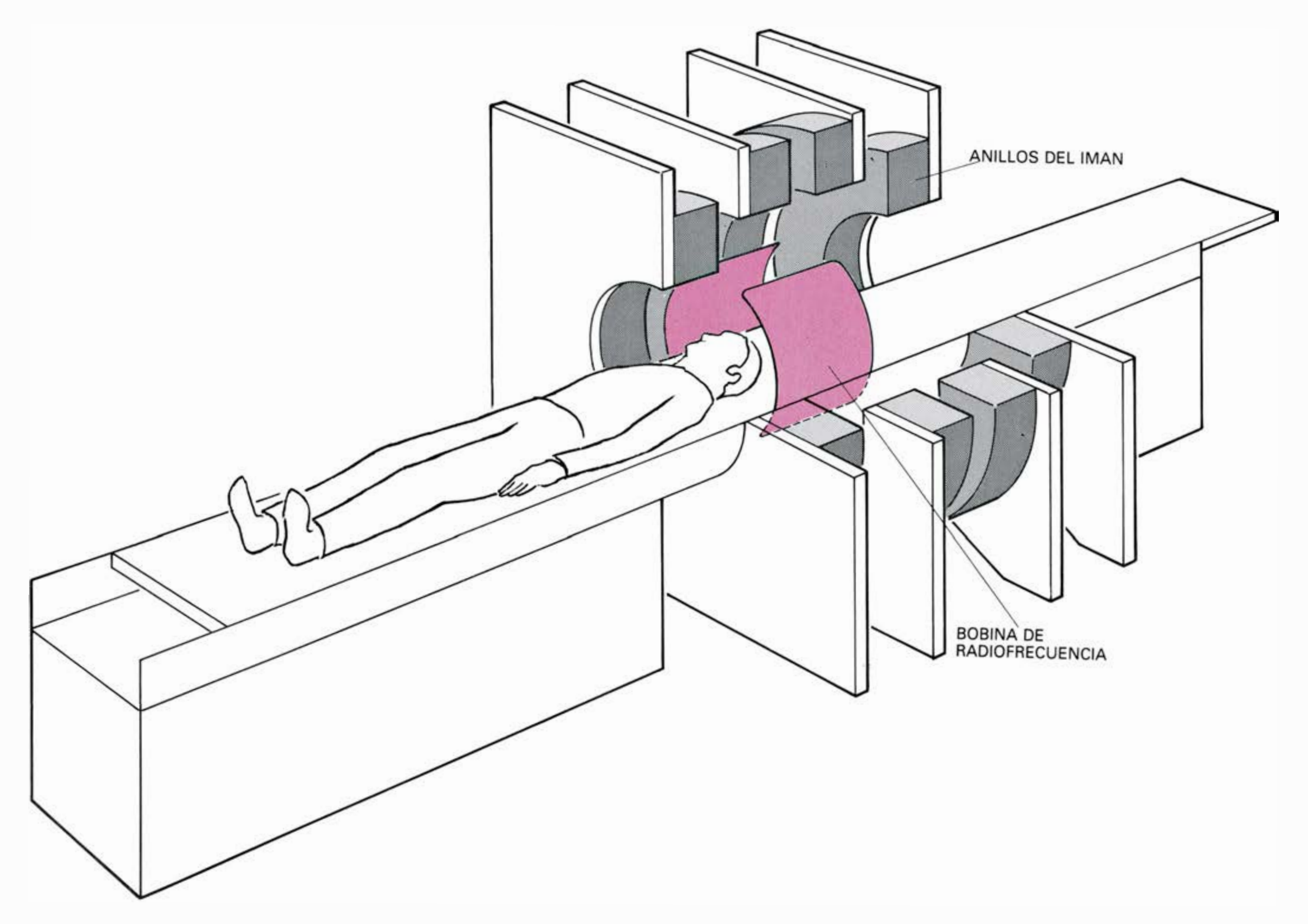
Debe subrayarse que la resolución espacial de una imagen de RMN no viene dictada por la longitud de onda de la radiación que la produce, como sucede en muchos sistemas de formación de imágenes. (La longitud de onda de la radiación de 15 megahertz en el espacio libre es de 20 metros.) En la formación de imágenes por RMN la resolución espacial depende de la uniformidad del campo magnético estático y de la fuerza de los campos de gradiente. El tiempo de relajación T^*_2 de la señal de inducción libre suele ser el proceso de amortiguación dominante en los actuales sistemas de formación de imágenes por RMN; por tanto, dicho valor fija una amplitud mínima de línea del espectro

de RMN. La fuerza de los gradientes de campo ha de alcanzar un valor suficiente para que los incrementos de la frecuencia en cada elemento (pixel a pixel) de la imagen final dominen la amplitud de la línea T^*_2. Algunos métodos de codificación espacial no exigen, sin embargo, tanto rigor en esta condición como otros.

La formación de imágenes por RMN en medicina se encuentra en su infancia, por lo que hay que acumular un volumen sistemático de datos antes de asegurar la eficacia de la técnica en la detección de lesiones en las distintas partes del cuerpo. Los primeros resultados clínicos obtenidos en el Hospital General de Massachusetts con dos sistemas de formación de imágenes fabricados por la Technicare Corporation sugieren que la RMN puede resultar particularmente idónea para identificar tejido necrosado (muerto), isquemias (anemia local provocada por obstrucción mecánica de la irrigación sanguínea), tumores y enfermedades degenerativas de múltiple naturaleza. El contraste del tejido blando es superior al de las técnicas de rayos X, y la sensibilidad para la detección de lesiones, en general, parece muy alta.

Un camino de gran futuro es el de la administración de "trazadores de RMN", o material de contraste, por vía oral o por inyección. Los iones paramagnéticos, como los del manganeso (Mn^{++}), tienen electrones no apareados, de forma que poseen un momento magnético y tienden a alinearse en el campo magnético estático de un sistema de RMN. Estos iones incrementan ligeramente el campo magnético local y modifican sustancialmente el tiempo de relajación de los núcleos sensibles a la RMN. La variación resultante en la intensidad de la imagen permitiría conocer la distribución de la especie paramagnética. Este tipo de trazadores podrían desempeñar un papel análogo al de los trazadores radiactivos administrados en medicina nuclear. En varios laboratorios se está investigando qué posibilidades encierra la resonancia magnética nuclear a la hora de determinar la velocidad del flujo sanguíneo a través de los vasos mayores. Los resultados de experimentos recientes demuestran asimismo la viabilidad de "congelar" la acción del corazón en diversas partes de su ciclo.

Quizá la mayor esperanza esté puesta en la formación de imágenes de núcleos que no sean hidrógeno, particularmente el núcleo del fósforo. El fósforo es un importante constituyente de las moléculas de alta energía, como el trifosfato de adenosina (ATP) y la fosfocreatina, que median la transferencia de energía en la célula viva. La espectroscopía por RMN del fósforo en determinados volúmenes del cuerpo revela varios picos de resonancia, químicamente desplazados, cuyas alturas corresponden a las concentraciones de cada compuesto de fósforo. Partiendo del conocimiento de estas concentraciones cabe deducir el estado metabólico de los órganos internos, y, andando el tiempo, quizá pudiera incorporarse esa facultad a un instrumento de formación de imágenes. Los próximos años prometen una importante mejora de la calidad de las imágenes por RMN, y una mayor diversidad de aplicaciones para la resonancia magnética nuclear en la práctica clínica.



IMAN PARA IMAGENES DE RMN, representado esquemáticamente. Los cuatro anillos sostienen las espiras que generan un campo magnético estático (B_0). El pulso de radiofrecuencia lo produce una bobina de paneles curvados. Las señales de RMN inducidas por los pulsos de radiofrecuencia las detecta la misma bobina, que las envía a un ordenador. Los sistemas más costosos, que generan campos magnéticos más potentes, utilizan imanes superconductores.

Ciencia y sociedad

Sucesión y evolución

Los ecólogos de ciertas escuelas muestran una especial afición por destacar procesos de cambio en los ecosistemas y asignarles flechas con sentidos bien definidos. Valga de ejemplo la sucesión. Se entiende por tal la lenta transformación en bosque de un campo abandonado, a través del incremento de la biomasa, normalmente con aumento de la complejidad y siempre con reducción de la energía intecambiada por unidad de tiempo y por cantidad de biomasa conservada. Otro ejemplo es el modelo de evolución gradual, que se aprecia en el registro fósil de varios grupos de organismos, como foraminíferos, ostrácodos, amonites, caballos y elefantes. Muchos reconocen en estas regularidades ciertas "leyes" formales que atañerían al aumento del tamaño, la especialización y la irreversibilidad.

Una representación completa de la naturaleza debería ser más amplia y, más allá de los procesos continuos de sucesión y evolución regular, habría de considerar los procesos, más o menos opuestos, que conducen a la destrucción o regresión de la cubierta vegetal. Habría de dar asimismo alguna explicación evolutiva de lo que ocurre cuando, y allí donde, haya de incluirse una componente importante de impredictibilidad e inventiva.

Seguir el sentido de una flecha que señala sucesión o evolución progresiva quizá resulte más atractivo, pues se trata de procesos lentos, hasta cierto punto predecibles, característicos del mundo orgánico y fáciles de racionalizar, sin duda por profundas razones psicológicas. Su sentido opuesto, el del cambio, se nos presenta impredecible, catastrofista y excesivamente lento o demasiado rápido para proporcionarnos un objeto de contemplación satisfactorio. Se nos ofrece una naturaleza surcada por procesos que siguen caminos alternativos; y sólo porque un sentido de cambio es rápido y el otro lento resulta más fácil atender al segundo, cuyas trayectorias ocupan un plano dominante en el escenario.

Detengámonos en algunos ejemplos. Así, el crecimiento de una población bacteriana en un frasco de cultivo. Comienza con un consumo de los nutrientes a su alcance rápido, derrochador y de pobre rendimiento. Remiten luego los procesos biológicos y bioquímicos y aparece, e irá aumentando, cierta diferenciación en el espacio y en el tiempo, sobre todo si no se agita el frasco y se trabaja con varios clones.

Un segundo ejemplo sería la vida en las deposiciones del ganado en un prado de montaña. No puede predecirse dónde vaya a caer la deposición; su utilización por parte de los organismos preexistentes en el excremento o en el suelo es rápida y diríamos que ineficiente, pero no cabe duda de que, a su tiempo, aparecerán los organismos adecuados, crecerán secuencialmente y desarrollarán un patrón muy regular de aprovechamiento, hasta que la inestabilidad local generada por la deposición se asimile al proceso de persistencia y sucesión del conjunto del prado.

Volvamos a la formación de un bosque. La destrucción de la vegetación arbórea anterior debióse a energía externa (huracán, excavadora) o a la acción, catastrófica, de energía acumulada en el ecosistema y desencadenada desde el exterior. Pensemos en un incendio forestal, cuyo equivalente, a escala de la biosfera, sería la ignición de los combustibles fósiles. La persistencia de un campo de cultivo requiere una explotación continua e impedir la invasión por malas hierbas y arbustos. Si se abandona, siempre habrá diásporas que acaben por reconstruir el bosque, dentro del cuadro de limitaciones que imponga el clima local.

¿Qué ocurre en el fitoplancton? La sucesión de un bosque tardá siglos; décadas dura la vida de un árbol. Pero el tiempo de generación de los organismos del fitoplancton va de menos de un día a unos pocos días; en unas cuantas semanas se desarrollan sucesiones equivalentes al proceso de reconstrucción de un bosque. El enfriamiento invernal o cualquier racha de viento mezclan el agua y uniformizan las distribuciones verticales de nutrientes y organismos germinales. La sucesión comienza con el dominio de los organismos pequeños y de rápida multiplicación. La turbulencia decae lentamente, se estratifica el agua y les llega su vez a los organismos mayores, que suelen estar dotados de motilidad, cuyo metabolismo de crecimiento disminuye a medida que consumen los nutrientes.

Las regularidades no se limitan a sistemas compuestos por organismos vivos. En cualquier acúmulo potente de sedimentos se advierten por lo general ciclos de sedimentación. Todos los ciclos empiezan con material grueso y, lentamente, de forma gradual, se pasa a material de grano más fino. La siguiente transición de material fino a grueso es abrupta. La racionalización del hecho guarda relación con otra que vale también para las secuencias de fitoplancton. Una entrada de energía desde fuera mezcla el agua o transporta partículas de materia sólida; la regresión de la turbulencia se refleja en una sucesión planctónica que se ajusta a una explicación ecológica en términos de competición bajo condiciones variables, u otras, en el caso de los sedimentos, en el descenso gradual del tamaño de las partículas en deposición. Caben explicaciones similares para otros procesos de diferenciación geológica.

En cada uno de los ejemplos precedentes se advierte un punto de inflexión entre un período de procesos rápidos, dirigidos desde el exterior e impredecibles desde dentro, y un período de cambio lento y controlado desde el interior, dependiente en gran parte de la organización ya existente. Una descripción lapidaria diría que el sistema construye la historia durante el primer período, para registrar luego parte de ella a lo largo del segundo. Es difícil individualizar las cosas en la naturaleza, tal es el solapamiento y la combinación, a diferentes escalas, de los múltiples procesos que se desarrollan a ritmo muy dispar. Valga de ejemplo recordar la sucesión en las deposiciones del ganado a medida que va asimilándose en la sucesión global del prado.

Los ecosistemas están sometidos a la aportación de energía, cuya intensidad y periodicidad pueden considerarse aleatorias. Cada entrada se filtra por toda la organización y genera una respuesta relativamente lenta y más bien larga, según de qué organización se trate. Las fluctuaciones de los ecosistemas, por tanto, no son simétricas; su representación nos da una configuración en diente de sierra. Lo que rebaja la utilidad del análisis del cambio del ecosistema mediante la suma de ondas sinusoidales. Si las entradas estocásticas son frecuentes, en relación con el tamaño y la longevidad de los organismos importantes, pueden sumarse con más facilidad en una salida regular. Una organización adecuada puede convertir en fuente de organización mayor lo que en un ecosistema organizado de otro modo representaría un hecho generador, en potencia, de una catástrofe. La transferencia de energía entre distintas escalas difiere bastante en un medio fluido de lo que ocurre en un sistema organizado. Las oscilaciones simétricas sólo podrían darse en sistemas compuestos enteramente por energía,

algo que se aparta de la realidad del mundo.

Los sistemas de magnitud suficiente pueden analizar las fluctuaciones pretéritas y reconocer ritmos en ellas. Si se mantiene la regularidad y pueden más o menos anticiparse las futuras entradas de energía, se habrá hallado la manera de reducir las asimetrías en las respuestas a las entradas. En cuanto la vida reconozca la melodía, lo inesperado dejará de serlo tanto. Los ritmos anticipan hechos y proporcionan una ventaja obvia en la competencia y selección. A un nivel más personal y psicológico quizás ayuden a explicar los efectos tranquilizantes de la música que no carece de ritmo.

El modelo de sucesión y cambio, más general, permite una descripción sumamente simple de los dos períodos. El primer estadio, de cambio catastrófico, seguido de la rápida colonización de un espacio ecológicamente vacío por toda clase de oportunistas, con un uso despilfarrador de los recursos, puede describirse como el resultado de una mezcla rápida y total de los reactivos en una situación que admite un incremento máximo y rápido de la entropía, algo así como un gas que se expande y se mezcla en un depósito. El segundo período, la sucesión convencional, se caracteriza por la diferenciación local y el incremento del número de elementos estructurales distinguibles. A continuación, cada uno de los subsistemas componentes que se han diferenciado se comporta como un sistema abierto y, según la teoría, la cantidad de energía realmente intercambiada, o el correspondiente incremento de la entropía, decrece constantemente por unidad de biomasa total conservada. Esta biomasa sería una medida de la información ganada por el sistema.

En el estado inicial no pueden señalarse distinciones. A medida que transcurre el tiempo decrece la tasa de recambio (turnover), aunque no en igual proporción en los distintos compartimentos, resultando de ello un enriquecimiento de la estructura. La información, u organización, se relaciona directamente con la posibilidad de distinguir divisiones en la biomasa, una de las condiciones de su diferenciación. El proceso histórico de diferenciación interna es esencial en la noción de sucesión, como reconocen los naturalistas. Pero la respetable sabiduría de los teóricos ha seguido otros derroteros y ha producido modelos matemáticos que pretenden enlazar las entradas y salidas de las piscifactorías, ha clarificado los desplazamientos de la estructura del ecosistema en relación con el grado de *stress* o de explotación a que se somete el sistema y con el desarrollo histórico del mismo. Los modelos al uso requieren aún de otro grado de complicación para acomodar las consecuencias del cambio histórico en una organización persistente, o las de su destrucción. Quizá no sea inoportuno afirmar que los modelos actuales lo son de "primera ley", y que lo que en realidad nos está haciendo falta son modelos de "segunda ley". De no darse este paso al frente, la mayoría de los ejercicios de modelación parecerán inútiles, recordándonos el aforismo lúdico de que, en ecología, cualquier expresión que sobrepase los 10 centímetros, es falsa. (Ramón Margalef.)

Deuterón diabólico

Según el modelo quark de las partículas subnucleares, ni el protón ni el neutrón son elementales: cada una de ellas consta de tres quarks. De donde se infiere que un núcleo atómico pesado contiene muchos quarks. Se supone que protón y neutrón mantienen, en el núcleo, su propia identidad singular, pero cabe la posibilidad de que los quarks se organicen de otras formas. Sverker Fredriksson y Magnus Jändel, del Real Instituto de Tecnología de Estocolmo, acaban de sugerir si no podrían considerarse algunos hechos como prueba de la existencia de una nueva clase de materia nuclear cuya organización interna presente un tal barajamiento de quarks. El nuevo estado de materia propuesto ofrece la misma composición en quarks que el deuterón, el núcleo del isótopo pesado de hidrógeno, constituido por un protón y un neutrón. Pero en vez de condensarse en dos tripletes, los seis quarks del hipotético núcleo se organizan en tres grupos de dos quarks. Fredriksson y Jändel denominan deuterón diabólico al nuevo núcleo y analizan sus hallazgos en *Physical Review Letters*.

La prueba principal en la que Fredriksson y Jändel fundan su hipótesis la aportaron, en el otoño de 1980, un grupo de experimentadores del Lawrence Berkeley Laboratory de la Universidad de California. Los investigadores dirigieron un haz de oxígeno ionizado o de átomos de hierro de alta energía sobre una emulsión fotográfica que luego se revelaba para poner de manifiesto las trazas de las partículas dotadas de carga eléctrica. Cada colisión de un átomo de hierro del haz con un núcleo de la emulsión origina una lluvia de partículas hijas; muchas de las partículas hijas colisionan a su vez, creando lluvias secundarias, y así sucesivamente. Después de revelar la emulsión, los experimentadores miden la distancia entre colisiones para cada partícula hija. La distancia media entre colisiones, o el recorrido libre medio de una partícula, sirve de parámetro para medir la intensidad de sus interacciones con otras partículas. Un recorrido libre medio menor indica que la partícula tiene una mayor tendencia a chocar con un núcleo; en muchos casos, la probabilidad de colisión tiene que ver con el tamaño de la partícula. Los investigadores hallaron que alrededor de un seis por ciento de las partículas creadas en la emulsión presentaban un recorrido libre medio extraordinariamente corto. Ningún proceso nuclear conocido justificaba esos resultados.

El deuterón diabólico, preconizado por Fredriksson y Jändel para explicar esos resultados, no es mayor que el deuterón ordinario; su corto recorrido libre medio viene determinado por su interacción eminentemente fuerte con otros núcleos. La interacción es muy parecida a la que hay entre átomos químicamente activos. Cuando los electrones de un átomo están unidos al núcleo atómico de un modo más laxo, la fuerza electromagnética que liga los electrones al núcleo se "escapa" fuera del núcleo atómico. El átomo tiene, por tanto, mayor tendencia a encontrarse con otros átomos y a formar enlaces químicos con ellos. Análogamente, en el deuterón diabólico los tres pares de quarks han de ocupar un estado excitado y así se encuentran menos intensamente ligados entre sí que lo que lo están el neutrón y el protón en el deuterón ordinario. En consecuencia, la fuerza que los mantiene unidos, que se denomina fuerza de color, puede traspasar la inmediata vecindad de los pares de quarks y actuar en un área mucho mayor. Fredriksson y Jändel han propuesto varios caminos que permitirían observar ese endiablado deuterón.

¿Son las enanas blancas progenitoras de las supernovas?

El origen de las explosiones de supernova y de los púlsares constituye uno de los problemas clave de la astrofísica actual. El descubrimiento de un púlsar (y, por tanto, de una estrella de neutrones) en la nebulosa del Cangrejo y la relación de ambos objetos con la explosión de supernova acaecida el año 1054 indican que esta última y el púlsar se produjeron simultáneamente.

Argumentos basados en las observaciones astronómicas y en la teoría de la evolución estelar sugieren que las supernovas de tipo I (las más energéticas,

carentes de rayas de hidrógeno en su espectro y presentes en todo tipo de galaxias) se deben a la explosión de enanas blancas de carbono-oxígeno. Por el contrario, las supernovas de tipo II (menos energéticas, ricas en hidrógeno y presentes sólo en los brazos de las galaxias espirales) deben ser el resultado de la explosión de estrellas cuya masa exceda las 10 masas solares.

Los problemas surgen cuando se intenta explicar el mecanismo responsable de tales explosiones. Una estrella masiva que se encuentre en las últimas fases de su vida colapsa con gran facilidad sobre sí misma y no origina explosión alguna. Por otra parte, una enana blanca de carbono-oxígeno explota fácilmente sin dejar ningún residuo ligado de las características de una estrella de neutrones. Simplificando mucho puede decirse que el problema consiste en encontrar mecanismos que permitan la explosión de estrellas masivas y otros que conlleven el colapso total o parcial de las enanas blancas.

Actualmente se piensa que las enanas blancas que forman parte de un sistema binario compacto están detrás de los fenómenos de nova. El mecanismo que las produce es sencillo: la estrella compañera transfiere materia rica en hidrógeno a la enana blanca de manera que sobre la superficie de esta última se va formando una capa de masa creciente. Cuando ésta alcanza un cierto valor crítico, el hidrógeno fusiona explosivamente para dar helio, manifestándose la energía liberada en la reacción en forma de nova.

Se ha intentado recurrir a un "guión" similar para explicar el origen de las supernovas de tipo I, aunque en este caso la explosión puede originarse tanto en la base del material recientemente incorporado como en el centro de la estrella, dependiendo ello del ritmo al que se efectúe la transferencia de materia; en cualquier caso, la explosión afecta al conjunto de la estrella. La debilidad de este tipo de modelos radica en que la catástrofe estelar no deja residuo ligado alguno en forma de estrella de neutrones.

Nuestro modelo para explicar las supernovas de tipo I y la formación de estrellas de neutrones transcurre en un escenario semejante al que acabamos de describir. La única diferencia consiste en que la enana blanca que consideramos es muy densa y fría, de forma que su interior se encuentra cristalizado. Estas condiciones pueden alcanzarse con relativa facilidad si la enana en cuestión forma parte de un sistema estelar doble en el que la compañera es una estrella de masa pequeña, pues en este caso la transferencia de materia de la una a la otra no empieza hasta que han transcurrido lapsos de tiempo superiores a los mil millones de años.

La cristalización del interior de la enana provoca esencialmentre tres efectos: *a*) prolonga el límite de estabilidad térmica del carbono hasta densidades muy elevadas, *b*) reduce la velocidad con que se propaga el "fuego" nuclear y *c*) probablemente origina una diferenciación química de la estrella. Para ilustrar las consecuencias de estos efectos sobre la evolución de la estrella hemos estudiado el comportamiento de enanas blancas sólidas con distintos grados de separación química cuando sufren acreción de materia.

La densidad central de la enana blanca aumenta gradualmente debido al aporte de masa procedente de la compañera. Puesto que la estrella ha solidificado, el carbono-12 puede fusionar consigo mismo sólo gracias a las vibraciones del estado fundamental de la red cristalina (se dice que estamos en régimen picnonuclear), por lo que el ritmo de las reacciones es muy lento y depende únicamente de la densidad del material. Conforme aumenta la densidad central, las reacciones picnonucleares se aceleran, de modo que cuando se alcanza el valor de 10^{10} gramos por centímetro cúbico el calor generado funde el material y las reacciones pasan a producirse en régimen termonuclear y adquieren entonces carácter explosivo.

Sin embargo, a las densidades involucradas, el frente de combustión no se propaga a velocidad supersónica (detonación), sino que se desplaza a velocidad subsónica (deflagración); el valor concreto de la velocidad depende del modo de transmisión del calor. Ya que el ritmo de las reacciones picnonucleares es extremadamente sensible a la densidad, la licuefacción y posterior ignición se producen sólo en una pequeña región del centro de la estrella, mientras que el resto permanece sólido. Por tanto, la velocidad de propagación del frente de combustión está determinada por la conductividad electrónica del sólido. En estas condiciones, las capturas electrónicas sobre las "cenizas" de la combustión ocasionan una caída de presión que provoca el colapso de la zona central de la estrella, aunque las incertidumbres sobre el valor de los ritmos de estas reacciones no permitan asegurar categóricamente que se forme un residuo neutronizado.

No obstante, la consecuencia más decisiva de la cristalización es la posible separación de los componentes químicos de la estrella. Es bien conocido en metalurgia que algunas mezclas de metales sólo pueden solidificar sin cambio de composición si los integrantes intervienen en unas determinadas proporciones (mezcla eutéctica). En caso contrario, los componentes solidifican por separado. Se ha señalado recientemente que un fenómeno similar podría producirse en el proceso de cristalización del material constitutivo de las enanas blancas de carbono-oxígeno. Estos dos elementos sólo cristalizarán conjuntamente si las proporciones de carbono y de oxígeno son del 66,8 por ciento y del 33,2 por ciento en número, respectivamente (composición eutéctica). De no ser así, el cristal que se forme estará constituido por el elemento que sea sobreabundante con respecto a la mezcla eutéctica. En el caso de las enanas blancas el oxígeno será el primero en cristalizar, puesto que su abundancia es superior al 33,2 por ciento. De esta forma, el oxígeno cristalizado se irá acumulando en el centro de la estrella, al tiempo que el carbono se verá desplazado hacia la superficie. El tamaño del núcleo de oxígeno depende del tiempo transcurrido desde el inicio de la cristalización hasta la reactivación de la enana blanca como consecuencia de la acreción. La importancia de este proceso radica en que el material más explosivo, el carbono-12, queda relegado a regiones menos "peligrosas", con el consiguiente aumento de la estabilidad de la estrella.

Si el tiempo disponible para el enfriamiento ha sido suficiente, la estructura final consistirá en un núcleo de oxígeno puro rodeado de un manto de carbono puro. En este caso los aumentos de densidad y temperatura, consecuencia del aumento de la masa de la estrella que resulta de la transferencia de materia, no llegan a provocar ni la ignición del oxígeno en el centro ni la del carbono en las capas intermedias, ya que el primero es poco "explosivo" y el segundo se encuentra suficientemente lejos del centro. Así pues, el incremento de la densidad central proseguirá hasta que se alcance el valor de 2×10^{10} gramos por centímetro cúbico, momento en que el oxígeno empieza a capturar electrones transformándose en carbono-16. Este último es tan buen combustible como el carbono-12, por lo que en cuanto su abundancia comienza a ser apreciable tiene lugar su ignición y, a consecuencia de ella, también la del oxígeno.

La situación es en apariencia semejante a la del caso completamente mezclado, pero ahora la densidad central es el doble y, además, el frente de combustión se propaga a través de un combustible peor (el ^{16}O), por lo

que está garantizada la formación de una estrella de neutrones como resultado de la pérdida de presión debida a las capturas electrónicas sobre los productos de la combustión.

Si el tiempo de enfriamiento no es lo bastante largo, la estructura de la estrella al inicio de la transferencia de materia consistirá en un núcleo de oxígeno de tamaño variable, según sea la duración exacta de dicho período, rodeado por un manto de carbono-oxígeno. En este caso, el carbono se encontrará mucho más cerca del centro, por lo que los incrementos de densidad y de temperatura debidos a la transferencia de materia sobre la enana blanca pueden provocar su ignición antes de que el centro de la estrella alcance la densidad de 2×10^{10} gramos por centímetro cúbico, umbral para el inicio de las capturas electrónicas sobre el oxígeno. Tendremos, pues, ignición del carbono-12 lejos del centro con la subsiguiente formación de dos frentes de combustión, uno que se propaga hacia el interior, a través del núcleo de oxígeno, y otro que se dirige hacia la superficie, a través del manto de carbono-oxígeno. Este último será responsable de la explosión y expulsión de las capas externas de la estrella, a densidad relativamente baja y bien provistas de combustible (el ^{12}C), mientras que en la zona central volvemos a encontrarnos con un proceso de combustión que se propaga lentamente y que se alimenta de un material poco explosivo a altas densidades, por lo que probablemente dicho núcleo, o al menos una parte de él, podrá colapsar y dar así lugar a una estrella de neutrones.

Vemos, pues, que la conexión entre supernovas y estrellas de neutrones, apuntada por la presencia de un pulsar en el residuo de la supernova del Cangrejo, queda, en lo que se refiere a las supernovas de tipo I, contemplada y explicada en el marco de la evolución de enanas blancas de carbono-oxígeno que forman parte de un sistema doble y que han dispuesto del tiempo suficiente para que su interior cristalice antes de verse reactivadas por transferencia de materia. Si bien una visión global y definitiva del problema requiere tener en cuenta una serie de aspectos que han sido dejados al margen de este trabajo, entre los que podemos destacar el estudio detallado del comportamiento térmico de las capas incorporadas a la enana blanca durante la fase de transferencia de materia y el de la etapa de evolución dinámica que sucederá a la desestabilización de la estrella. (Ramón Canal, Jordi Isern y Javier Labay.)

Bases genéticas de la diversidad de anticuerpos

Durante el desarrollo de las células del sistema inmunitario, segmentos de ADN y ARN se mezclan y unen de múltiples maneras. Las combinaciones generan información para fabricar miles de millones de anticuerpos distintos

Philip Leder

El sistema inmunitario de un animal vertebrado tiene una capacidad prácticamente ilimitada para fabricar anticuerpos diferentes, que reconozcan y se unan a muchos millones de antígenos potenciales, o moléculas "ajenas". Ello quiere decir que las células del sistema inmunitario sintetizan muchos millones de tipos de moléculas de anticuerpos. Un anticuerpo es un conjunto de cadenas proteínicas, cadenas cuya estructura está determinada por una unidad de información genética: un gen. Por tanto, podría parecer que hubiera muchos millones de genes de anticuerpos. Sin embargo, el genoma de un mamífero, o complemento genético total, encierra a lo sumo un millón de genes, de los cuales sólo una pequeña fracción determinan anticuerpos. Tal paradoja, un número limitado de genes frente a una, diríase, ilimitada capacidad para fabricar anticuerpos diferentes, ha intrigado, desde hace más de dos décadas, a inmunólogos y genetistas. Empieza ahora a vislumbrarse la solución del rompecabezas, en lo que supone una estrecha colaboración de ambas disciplinas, inmunología y genética, y el triunfo inmediato de algunas de las nuevas técnicas de la genética molecular.

La respuesta, a grandes rasgos, es la siguiente: los genes que determinan, en última instancia, la estructura de cada anticuerpo no están presentes como tales en las células germinales (los espermatozoides en los machos y el óvulo en las hembras) ni en las células del embrión recién formado. Estas células, en vez de llevar una serie de genes completos y activos para fabricar anticuerpos, contienen trozos y piezas de los genes: un lote de componentes. Los componentes se mezclan en las células del sistema inmunitario denominadas linfocitos *B*, a medida que éstas se desarrollan y maduran. La mezcla puede dar lugar a un resultado diferente en cada uno de los millones de líneas celulares. Las distintas mutaciones amplían la diversidad. El resultado final es que, en los descendientes maduros de cada línea, se "monta" un único gen, cuya información se expresa en forma de un solo anticuerpo.

Proteínas y genes

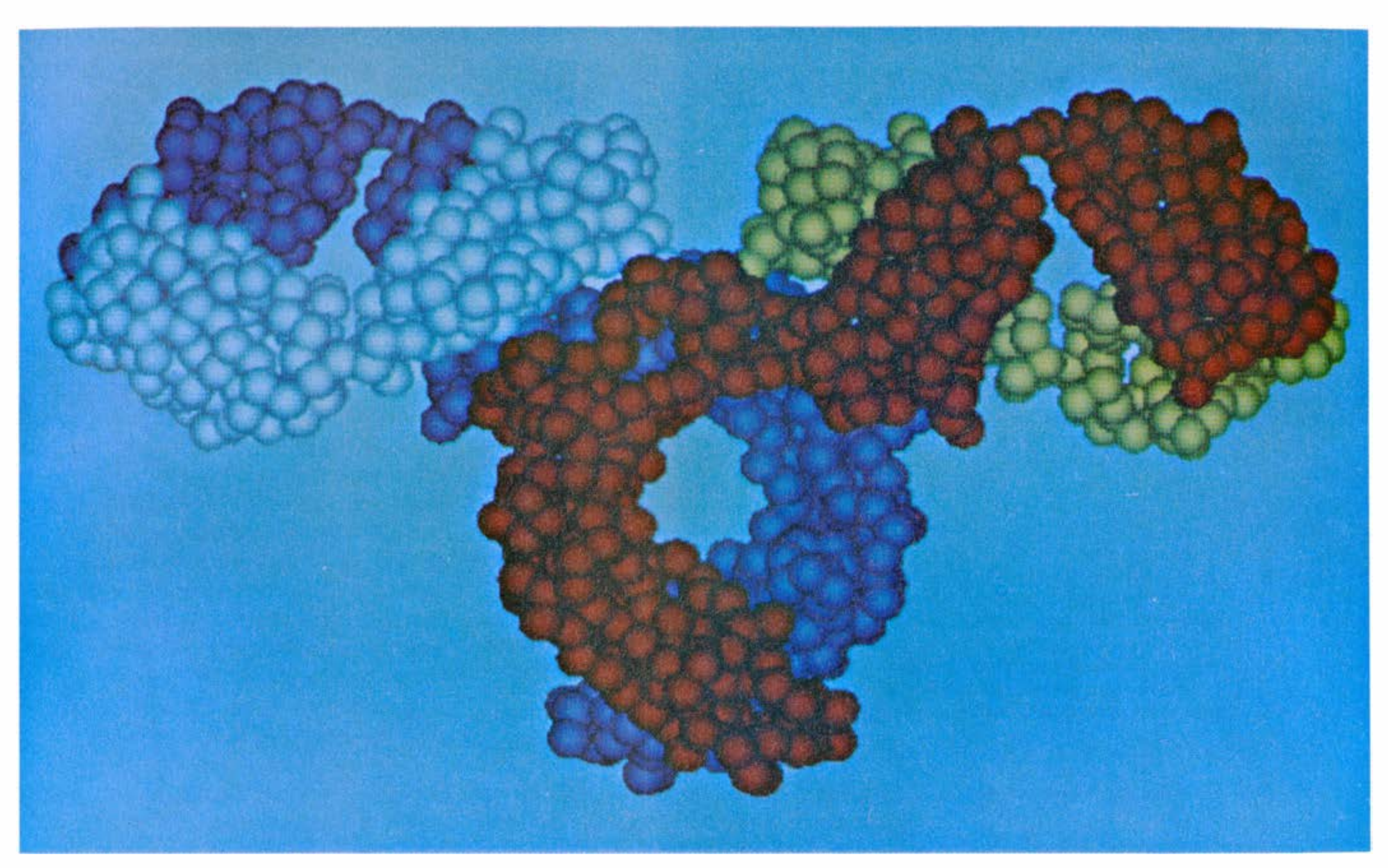
Como las demás proteínas, los anticuerpos están formados por subunidades llamadas aminoácidos. Hay 20 tipos de aminoácidos, que pueden unirse en cualquier combinación y constituir una cadena proteica. La composición de aminoácidos de una cadena y la secuencia en que éstos se disponen a lo largo de ella determinan la configuración tridimensional de ésta y la forma en que puede combinarse con otras. La sustitución de un aminoácido por otro, así como la alteración de la secuencia, cambian las propiedades de la proteína.

La composición y secuencia de aminoácidos de una cadena proteínica vienen determinadas por un gen, que es un segmento de otro tipo de cadena: una molécula de ácido desoxirribonucleico (ADN), el material hereditario de la célula. Las subunidades del ADN son cuatro nucleótidos, cada uno de los cuales se caracteriza por una de las cuatro bases químicas: adenina (A), guanina (G), timina (T) y citosina (C). Los nucleótidos se unen y forman una cadena de ADN; la secuencia de sus bases a lo largo de la cadena define la información que lleva el gen. La información se descifra según un código genético, cuyas palabras clave (denominadas codones) son tripletes de bases: CTG, por ejemplo, o AGC. En general, cada codón determina un aminoácido; una larga serie de codones suministra instrucciones para formar una cadena proteica completa, de cientos de aminoácidos. La información que porta el ADN, sin embargo, no se traduce directamente en una proteína. Primero se transcribe en una cadena de un ácido nucleico similar, el ácido ribonucleico (ARN). La nueva molécula informativa, denominada ARN mensajero, se traduce entonces en proteína.

En los organismos superiores, la información genética para la mayoría de las proteínas no está ordenada en una secuencia continua de codones de ADN. La mayoría de los genes se encuentran fragmentados: entre segmentos de secuencias informativas hay intercaladas secuencias no informativas. Un gen fragmentado se transcribe entero en ARN, incluidos los segmentos no informativos. A continuación, ese ARN "primario" se procesa, o madura: se eliminan las secuencias intercaladas y las secuencias informativas se unen para formar un ARN mensajero coherente.

Una molécula bifuncional

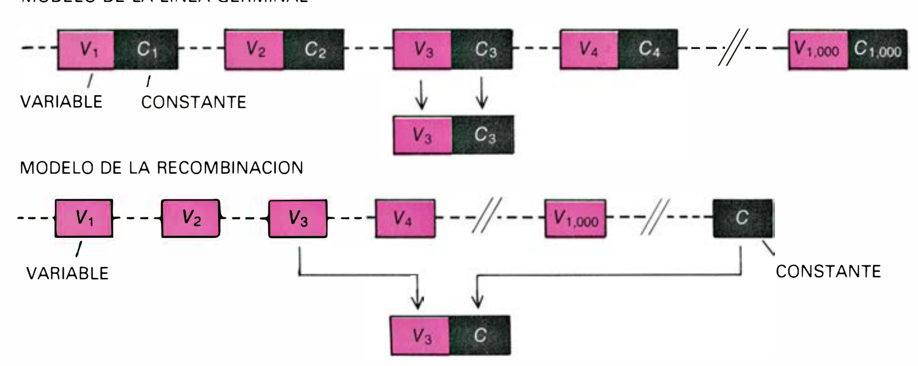
Las inmunoglobulinas, las moléculas de anticuerpos, poseen características estructurales que reflejan su función. Una molécula de anticuerpo consta de dos tipos de cadenas proteínicas relacionadas, denominadas ligera y pesada. Cuando, hace casi dos décadas, Norbert Hilschmann, a la sazón en el Instituto Rockefeller de Investigaciones Médicas, comparó las secuencias de aminoácidos de las cadenas ligeras de varios anticuerpos, encontró en ellas una propiedad particular. Las secuencias de las cadenas de distintos anticuerpos diferían unas de otras, pero las diferencias se limitaban a la primera mitad de cada cadena. El resto de la cadena tenía básicamente la misma se-



MOLECULA DE UN ANTICUERPO, conjunto de cuatro cadenas proteicas que se pliegan y entrecruzan formando una *T* partida. Se distinguen dos cadenas pesadas idénticas (*rojo y azul oscuro*) y dos cadenas ligeras igualmente idénticas (*amarillo y celeste*). En este modelo, construido por Richard J. Feldmann, del Instituto Nacional de la Salud, con la ayuda de un computador, las esferas representan aminoácidos, las subunidades de las cadenas proteicas.

CADENA PESADA
REGIONES HIPERVARIABLES
PUENTE DISULFURO
CADENA LIGERA
REGION VARIABLE
REGION CONSTANTE

LA NATURALEZA BIFUNCIONAL de un anticuerpo viene reflejada en su estructura, como se aprecia en este diagrama esquemático del modelo de molécula mostrado arriba. Cada cadena proteínica tiene una región variable y otra constante. La secuencia de aminoácidos de las regiones variables es peculiar de cada anticuerpo, mientras que las regiones constantes son las mismas en todos los anticuerpos de un tipo determinado. Las regiones variables reconocen y se unen a un antígeno específico; las constantes realizan tareas inmunológicas. Las cadenas se pliegan de suerte que sus regiones hipervariables, donde la secuencia de aminoácidos es especialmente voluble, se reúnen formando un sitio de reconocimiento del antígeno altamente específico.



DISPOSICION DE LOS GENES que determinan la cadena proteínica de un anticuerpo, según dos modelos diferentes. El modelo de la línea germinal propone que, en las células germinales (óvulos y espermatozoides), unos 1000 genes, cada uno de ellos con una secuencia de ADN que determina la región variable (*V*) de la cadena y otra secuencia con información para la región constante (*C*), codifican 1000 cadenas proteínicas distintas. En cada uno de los linfocitos productores de anticuerpos se expresaría uno de esos genes. Combinando 1000 genes de cadena pesada con otro millar de cadena ligera se podrían generar un millón de moléculas distintas de anticuerpos. Es difícil comprender, sin embargo, cómo habrían podido conservarse idénticas las 1000 secuencias de región constante en el transcurso de la evolución, mientras que en las secuencias de región variable se admitían cambios. El modelo de recombinación, propuesto en 1965 por William J. Dreyer y J. Claude Bennett, permite determinar el mismo número de cadenas proteínicas en las células germinales utilizando 1000 genes diferentes para la región variable, situados a cierta distancia de un único gen para la región constante. Cada uno de los genes para región variable podría recombinarse entonces con el gen *C*; todas las cadenas tendrían, pues, la misma región constante.

cuencia en todos los anticuerpos de un tipo determinado.

La presencia simultánea de variación y constancia en una sola molécula de proteína encierra gran significación funcional. De hecho, un anticuerpo es una molécula bifuncional; cada cadena tiene una región variable (aproximadamente la mitad de una cadena ligera y una cuarta parte de una pesada) y una región constante. Las regiones variables de las cadenas son las que se pliegan en el espacio para formar un lugar de unión antígeno-anticuerpo (el sitio que se une al antígeno particular contra el que va dirigido el anticuerpo). Cambiando la secuencia de aminoácidos de la región variable se modifica la estructura química del sitio de unión y se modifica, por tanto, la afinidad del anticuerpo por un antígeno, de la misma forma que alterando una muesca en la parte dentada de una llave ésta puede encajar en una cerradura diferente.

La región constante de las cadenas ligeras y pesadas de un anticuerpo es semejante al asidero de una llave, que es idéntico entre las llaves de un determinado tipo y desempeña una función común en todas ellas. La región constante de un tipo particular de moléculas de anticuerpo cumple idéntica misión en cada una de las moléculas. Hay, por ejemplo, dos tipos de cadenas ligeras en la mayoría de los animales vertebrados: kappa y lambda; todas las moléculas de anticuerpo deben tener cadenas ligeras de un tipo o del otro. En cualquier especie, la región constante de las cadenas ligeras kappa o lambda es idéntica a la región constante de otras cadenas del mismo tipo.

Todo anticuerpo tiene uno de los siguientes cinco tipos de cadenas pesadas: mu, delta, gamma, épsilon o alfa. El tipo de cadena pesada define la clase de inmunoglobulina, a saber, IgM, IgD, IgG, IgE o IgA. En los anticuerpos excretados que son de la clase IgM, por ejemplo, todas las cadenas pesadas tienen la misma secuencia constante mu y, todas las cadenas ligeras, la misma secuencia constante kappa o lambda. Las regiones variables difieren de un anticuerpo a otro, reflejando sus distintas especificidades antigénicas.

La región constante de las cadenas pesadas determina la función efectora de un anticuerpo, es decir, cómo realiza su tarea inmunológica en el organismo. Supongamos un anticuerpo cuya región variable es específica de un antígeno localizado en el polen de la ambrosía. Si la cadena pesada es del tipo delta, el anticuerpo IgD que se forma permanece asociado con la superficie de la célula que lo fabrica. Si está presente la cadena gamma, es probable que el anticuerpo IgG resultante circule por la sangre. Si se trata de la cadena épsilon, el anticuerpo (IgE) puede unirse a la superficie de una célula especializada que libera histamina, dando lugar a los síntomas de fiebre del heno o asma cuando el anticuerpo interacciona con el antígeno de la ambrosía. Todos los anticuerpos son específicos del mismo antígeno, es decir, el de la ambrosía. Además, esas mismas funciones efectoras se encuentran en anticuerpos dirigidos contra otros antígenos; la función efectora es independiente de la región variable.

Hipótesis de Dreyer-Bennett

Las características estructurales de los anticuerpos ofrecieron los primeros indicios sobre el origen genético de su diversidad. William J. Dreyer y J. Claude Bennett, que trabajaban en el Instituto de Tecnología de California a mediados de la década de los años 60, razonaron que combinando 1000 cadenas ligeras diferentes con 1000 cadenas pesadas distintas podrían generarse un millón de anticuerpos diferentes. No obstante, Dreyer y Bennett se preguntaban cómo podría organizarse la información genética que generase tal diversidad de proteínas. En particular, ¿qué organización daría cuenta de esa extraña estructura constante-variable?

Consideremos solamente el problema de un tipo de cadena ligera. Puede haber perfectamente 1000 genes, cada uno con información para una de las 1000 cadenas ligeras. Supóngase que así fuera. En el curso de la evolución operaría algún mecanismo que mantuviera inalterada la secuencia de la región constante de cada uno de los mil genes, mientras que la otra mitad de cada gen (la mitad que determina la región variable) disfrutaba de un amplio margen de mutación. Un mecanismo semejante parece improbable, ya que no hay motivos, desde un punto de vista biológico, para conservar la secuencia de aminoácidos de una determinada región constante de ésas. Las comparaciones entre inmunoglobulinas de diferentes individuos y especies revelan la insuficiente presión evolutiva para mantener identidad absoluta entre las secuencias de las regiones constantes; un cambio de unos pocos aminoácidos en la región constante no parece tener efectos perjudiciales.

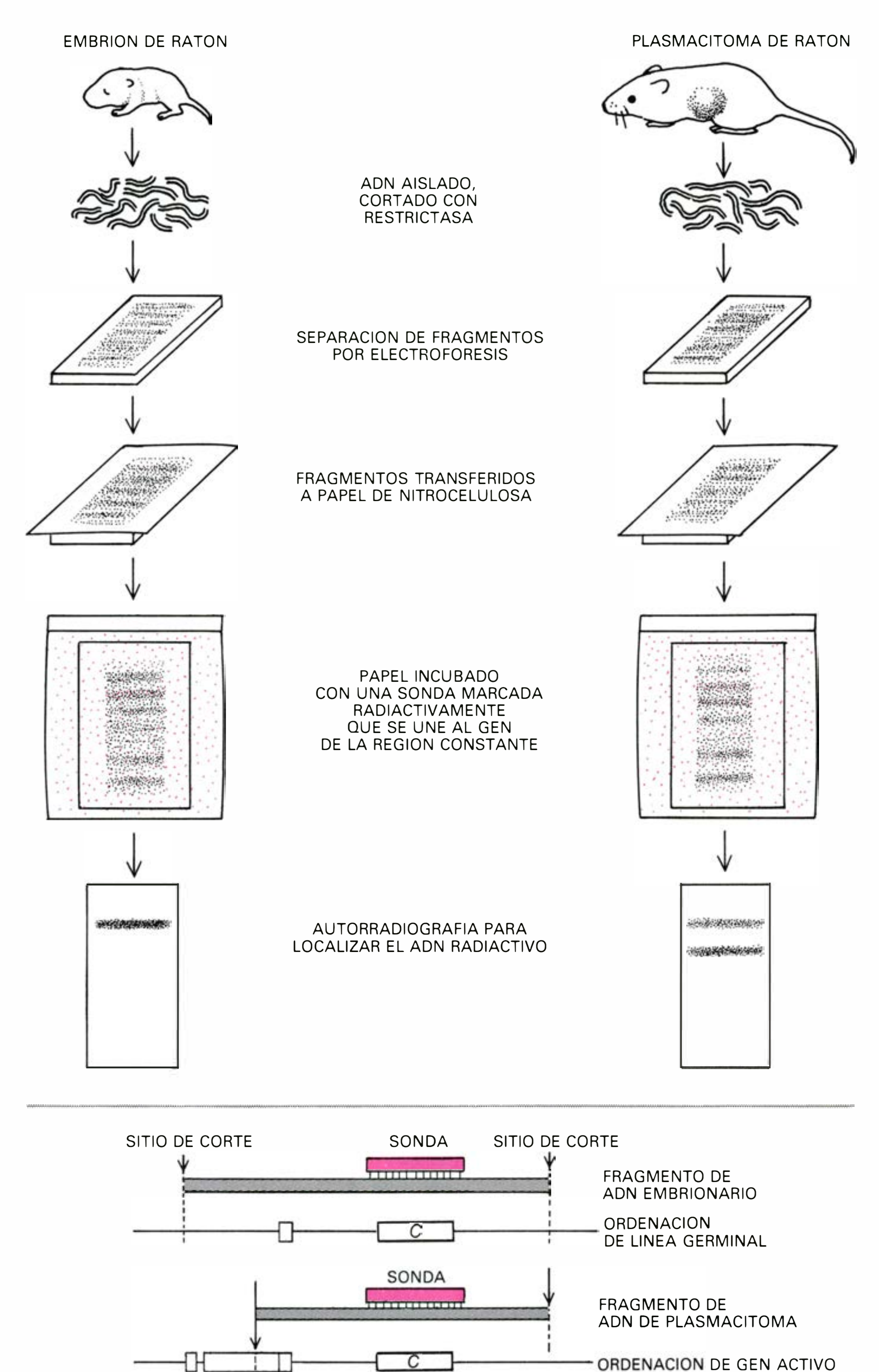
Dreyer y Bennett hicieron una propuesta radical. En vez de asumir que la información genética para la cadena ligera de un anticuerpo venía determinada por una serie continua de codones, propusieron que esa cadena la determinarían dos segmentos discontinuos de ADN, uno para la región variable y otro para la constante. Además postularon la existencia, en el ADN de las células germinales, de varios cientos o miles de genes separados para las regiones variables, pero sólo un gen para la región constante.

Suponiendo un único gen para la región constante, la propuesta de Dreyer y Bennett explicaba la aparición de una secuencia esencialmente igual en todas las regiones constantes de un tipo concreto de una especie dada. Si hay un solo gen para la región constante, cualquier mutación que se produzca en él debía alterar inmediatamente la secuencia de aminoácidos de todas las cadenas ligeras. En la propuesta estaba también implícita la idea de que la información presente en los elementos genéticos separados debía unirse de alguna forma para componer un mensaje genético contiguo y coherente y, por tanto, una sola proteína. La propuesta de Dreyer y Bennett atrajo inicialmente numerosas críticas; implicaba la existencia de genes fragmentados y mecanismos para unirlos, hechos que no tenían precedentes hasta entonces. Pero la idea se demostró correcta en sus líneas generales.

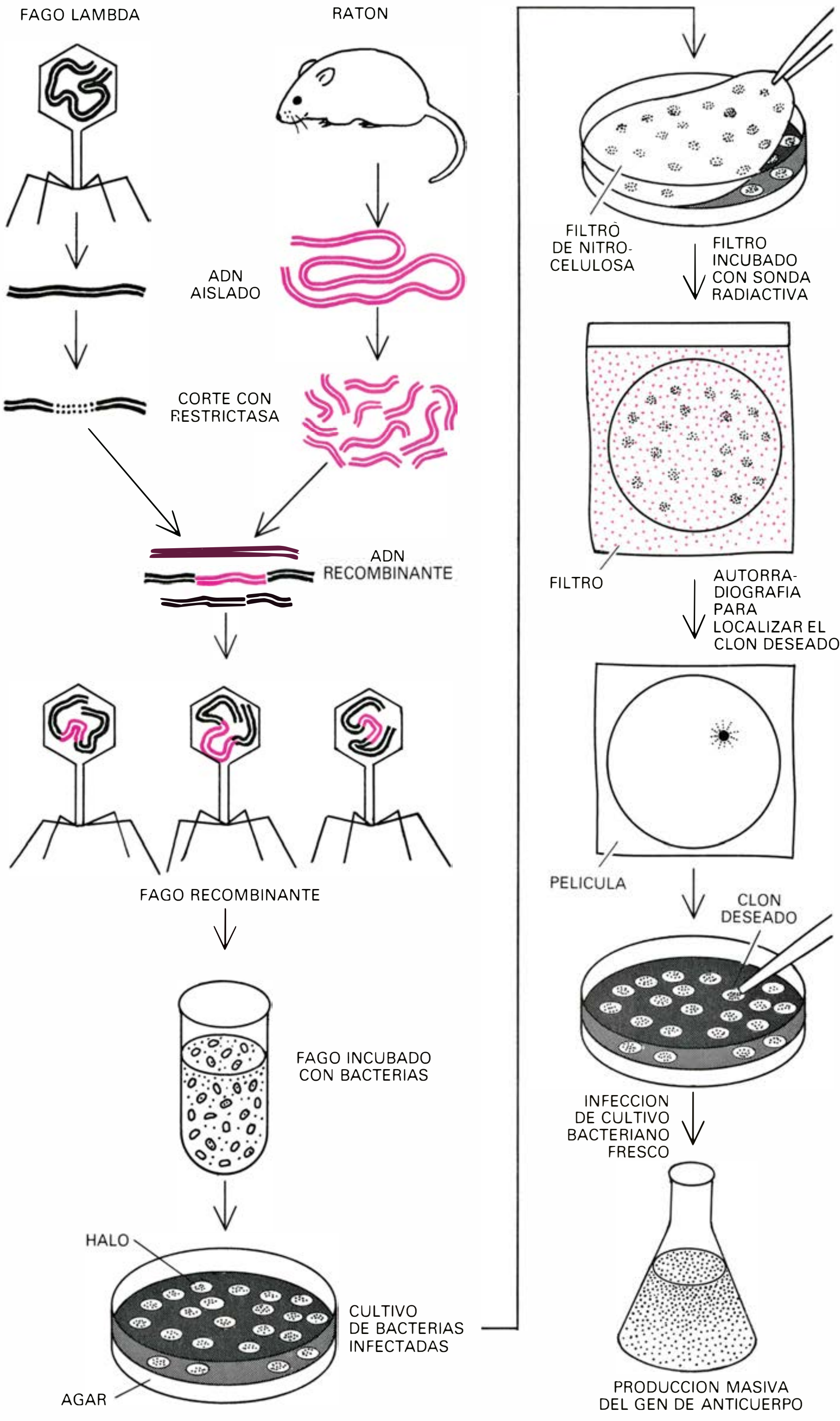
Primeros experimentos

En 1971 me dispuse a comprobar la hipótesis Dreyer-Bennett con mi colega David C. Swan, del Instituto Nacional de Salud Infantil y Desarrollo Humano. Nuestra estrategia era detectar y aislar el producto inicial de un gen de anticuerpo: su ARN mensajero. Una cadena de ADN copiada artificialmente a partir del mensajero (mediante el enzima transcriptasa inversa) serviría de sonda con la que detectar los genes de los anticuerpos en las células embrionarias y estimar su número. Ello nos permitiría comprobar la predicción central de la hipótesis de Dreyer-Bennett: que sólo hay un gen para la región constante o, como mucho, unos pocos, en contraposición a los muchos genes requeridos por modelos genéticos más sencillos.

Utilizamos la técnica de la cinética de hibridación. Dos segmentos de ácido nucleico cuyas secuencias de nucleótidos sean complementarias pueden hibridar, es decir, unirse entre sí. Si una de las secuencias se ha marcado con átomos de un isótopo radiactivo, las moléculas híbridas pueden identificarse por su radiactividad. La velocidad a la que una muestra de ADN radiactivo encuentra y se hibrida con cualquier molécula de ADN complementario (la cinética de hibridación) mide indirecta, aunque eficazmente, el número de dichas moléculas complementarias presentes en la preparación. Medimos a qué velocidad nuestra muestra (ADN copiado a partir del ARN responsable



REORDENACION DE LOS GENES durante el desarrollo. Se demuestra comparando fragmentos de ADN que contienen un gen de la región constante procedentes de embrión de ratón con los obtenidos de un plasmacitoma: tumor formado por una línea celular productora de anticuerpos. El ADN se digiere con una restrictasa, enzima que corta ese ácido en sitios específicos, caracterizados por presentar cierta secuencia nucleotídica. Los fragmentos resultantes se separan por tamaños mediante electroforesis en gel de agarosa y posteriormente se transfieren a papel de nitrocelulosa. Los fragmentos que contienen el gen de la región constante (*C*) se identifican incubando el papel con una sonda marcada radiactivamente (*color*), que podría muy bien ser una copia de ADN clonado obtenido a partir del ARN mensajero específico para la región constante de la cadena ligera. La sonda se hibrida con (se une a) cualquier ADN embrionario o de plasmacitoma que posea una secuencia de nucleótidos complementaria suficientemente larga. El ADN híbrido se identifica mediante autorradiografía, técnica en la que la radiación de la molécula marcada impresiona una película fotográfica. El ADN embrionario da lugar a una banda radiactiva, que corresponde a un único fragmento portador del gen *C* en la configuración de línea germinal. El ADN del plasmacitoma origina dos bandas. Representa una la configuración de la línea germinal, que se mantiene en uno de los dos alelos presentes en la célula productora de anticuerpos. La otra corresponde a un fragmento que avanza más en el gel, y es por tanto más corto. Representa al otro alelo, que ha sufrido reordenación y ahora lleva la secuencia *C* en la configuración de un gen activo de anticuerpo. La reordenación (*abajo*) ha eliminado uno de los sitios de corte y ha producido un sitio nuevo, más cerca de *C*.



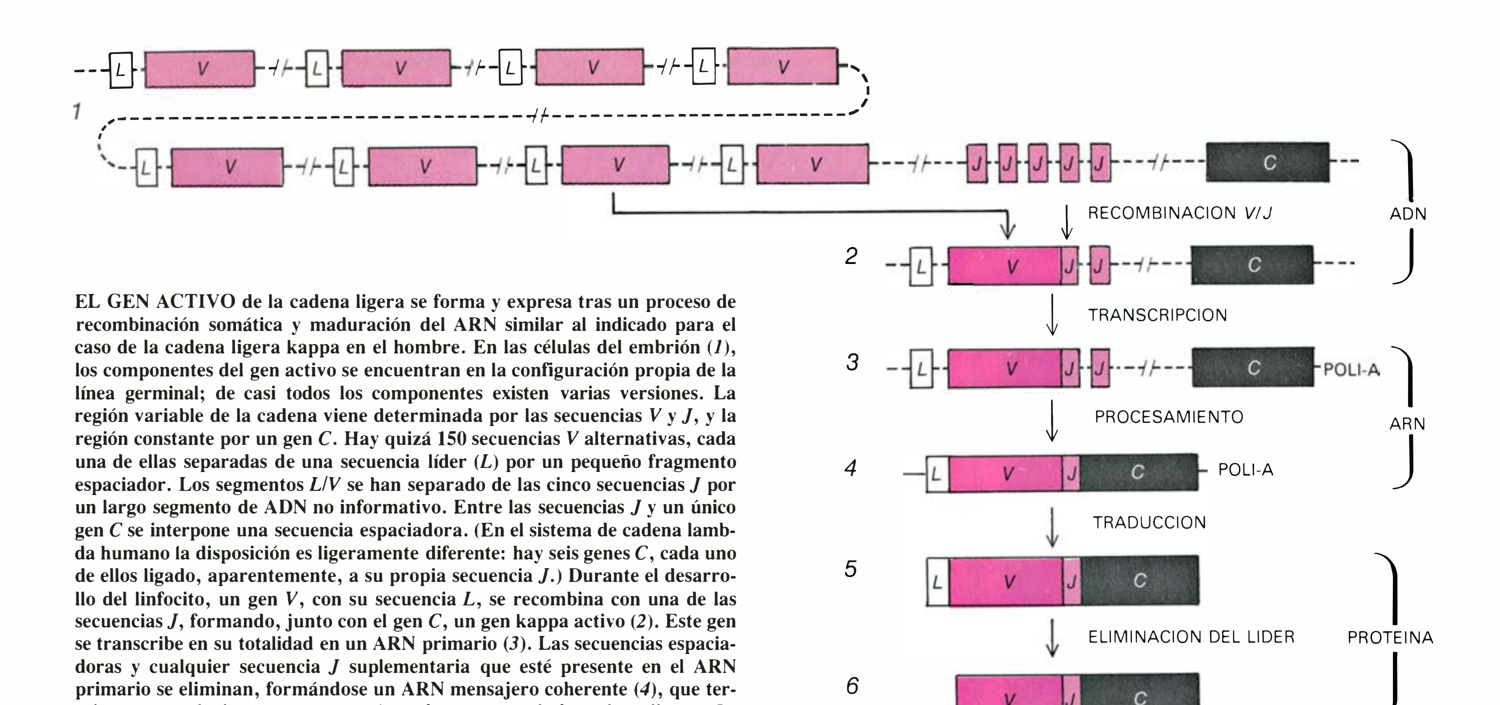
CLONACION DE UN GEN DE ANTICUERPO mediante la inserción de fragmentos de ADN de ratón en un virus bacteriano (fago lambda), posterior aislamiento de un clon vírico cuyo ADN contenga el gen y cultivo de ese virus para producir grandes cantidades del gen. Primero se digiere el ADN del fago con una restrictasa, que corta al ADN en tres fragmentos. El fragmento central no es vital para el fago. Puede sustituirse con un fragmento de ADN de ratón que se haya cortado con la misma restrictasa, de forma que los fragmentos del fago y del ratón tengan extremos complementarios ("extremos cohesivos") que pueden unirse. Los ADN recombinantes se empaquetan en nuevas partículas de fago, que se incuban posteriormente con *Escherichia coli*, la bacteria que infectan. Las bacterias infectadas se extienden sobre un medio de cultivo con agar. Los fagos proliferan matando a sus huéspedes e infectando a las células vecinas, lo que da lugar a la formación de áreas claras, denominadas halos, sobre la capa bacteriana; cada halo representa un clon de fagos. Los halos se transfieren a un filtro de nitrocelulosa y se eliminan las proteínas del fago, quedando retenido el ADN recombinante. Se incuba el filtro con una sonda, una copia de ADN fabricado a partir del ARN mensajero del gen deseado, marcada radiactivamente (*color*). La sonda se hibrida con cualquier ADN recombinante que posea una secuencia de ADN complementaria; mediante autorradiografía se detecta la posición del clon, que puede entonces seleccionarse del medio de cultivo y transferirse a un cultivo fresco del huésped bacteriano, que rinde grandes cantidades del anticuerpo.

de la región constante de la cadena ligera de ratón) se hibridaba con ADN representativo del genoma completo de embrión de ratón. Los resultados obtenidos, con la ayuda de Tasuku Konjo, que se nos juntó procedente de la Universidad de Kyoto, indicaban claramente que eran muy pocas las copias del gen de la región constante de la cadena ligera, no más de dos en cada célula quizá. En otros laboratorios se llegó pronto a conclusiones similares. La hipótesis de Dreyer-Bennett merecía tomarse en serio.

Si no hay más que unas cuantas copias del gen de la región constante y muchos genes de la región variable, tiene que existir algún mecanismo encargado de unir la información en una secuencia coherente. La forma más económica para realizar esa tarea sería a nivel de gen, esto es, unir dos secuencias de ADN, que están separadas en una célula embrionaria, para formar una única secuencia activa en el núcleo de un linfocito maduro y productor de anticuerpos. Tal reorganización del ADN durante la diferenciación y desarrollo de las células somáticas, o corporales, se denomina recombinación somática. El descubrimiento de los enzimas denominados restrictasas, que cortan el ADN en sitios específicos, permitió un importante experimento en apoyo de la idea de recombinación somática.

Susumu Tonegawa y Nobumichi Hozumi, del Instituto de Inmunología de Basilea, compararon ADN de embrión de ratón con ADN de un plasmacitoma: un tumor formado por células plasmáticas, o linfocitos *B* maduros, que producen moléculas de anticuerpo de un solo tipo y especificidad. Las células de esos tumores, la mayor parte de ellas preparadas por Michael Potter, del Instituto Nacional del Cáncer, proveen al investigador de grandes cantidades de un anticuerpo particular y del ADN y ARN que lo determinan; estas células han sido de importancia inestimable en muchas de las líneas de investigación inmunológica.

Tonegawa y Hozumi razonaron que si el ADN que rodeaba a los genes de los anticuerpos había sufrido recombinación somática en las células tumorígenas productoras de anticuerpos, su organización sería diferente de la del ADN de las células embrionarias. La diferencia se reflejaría en la organización de los sitios que corta una determinada restrictasa. Por ejemplo, si el gen de la región constante en el ADN embrionario estaba rodeado por dos sitios



EL GEN ACTIVO de la cadena ligera se forma y expresa tras un proceso de recombinación somática y maduración del ARN similar al indicado para el caso de la cadena ligera kappa en el hombre. En las células del embrión (*1*), los componentes del gen activo se encuentran en la configuración propia de la línea germinal; de casi todos los componentes existen varias versiones. La región variable de la cadena viene determinada por las secuencias *V* y *J*, y la región constante por un gen *C*. Hay quizá 150 secuencias *V* alternativas, cada una de ellas separadas de una secuencia líder (*L*) por un pequeño fragmento espaciador. Los segmentos *L*/*V* se han separado de las cinco secuencias *J* por un largo segmento de ADN no informativo. Entre las secuencias *J* y un único gen *C* se interpone una secuencia espaciadora. (En el sistema de cadena lambda humano la disposición es ligeramente diferente: hay seis genes *C*, cada uno de ellos ligado, aparentemente, a su propia secuencia *J*.) Durante el desarrollo del linfocito, un gen *V*, con su secuencia *L*, se recombina con una de las secuencias *J*, formando, junto con el gen *C*, un gen kappa activo (*2*). Este gen se transcribe en su totalidad en un ARN primario (*3*). Las secuencias espaciadoras y cualquier secuencia *J* suplementaria que esté presente en el ARN primario se eliminan, formándose un ARN mensajero coherente (*4*), que termina por traducirse en una prøteína: el precursor de la cadena ligera (*5*). La cadena madura pierde el líder (*6*) al atravesar la membrana celular.

de restricción (sitios por donde corta una restrictasa particular), separados entre sí por 5000 nucleótidos, tras la rotura el gen estará incluido en un fragmento de ADN de 5000 nucleótidos de largo. Si la recombinación somática eliminaba uno de los sitios y producía otro nuevo, el ADN de plasmacitoma daría lugar a un fragmento de la región constante mayor o menor de 5000 nucleótidos. Tonegawa y Hozumi demostraron que la organización de los genes de la cadena ligera no era la misma en las células embrionarias que en las células productoras de anticuerpos. La activación de los genes durante el desarrollo se acompañaba de recombinación somática: los genes se entremezclaban.

Clonación

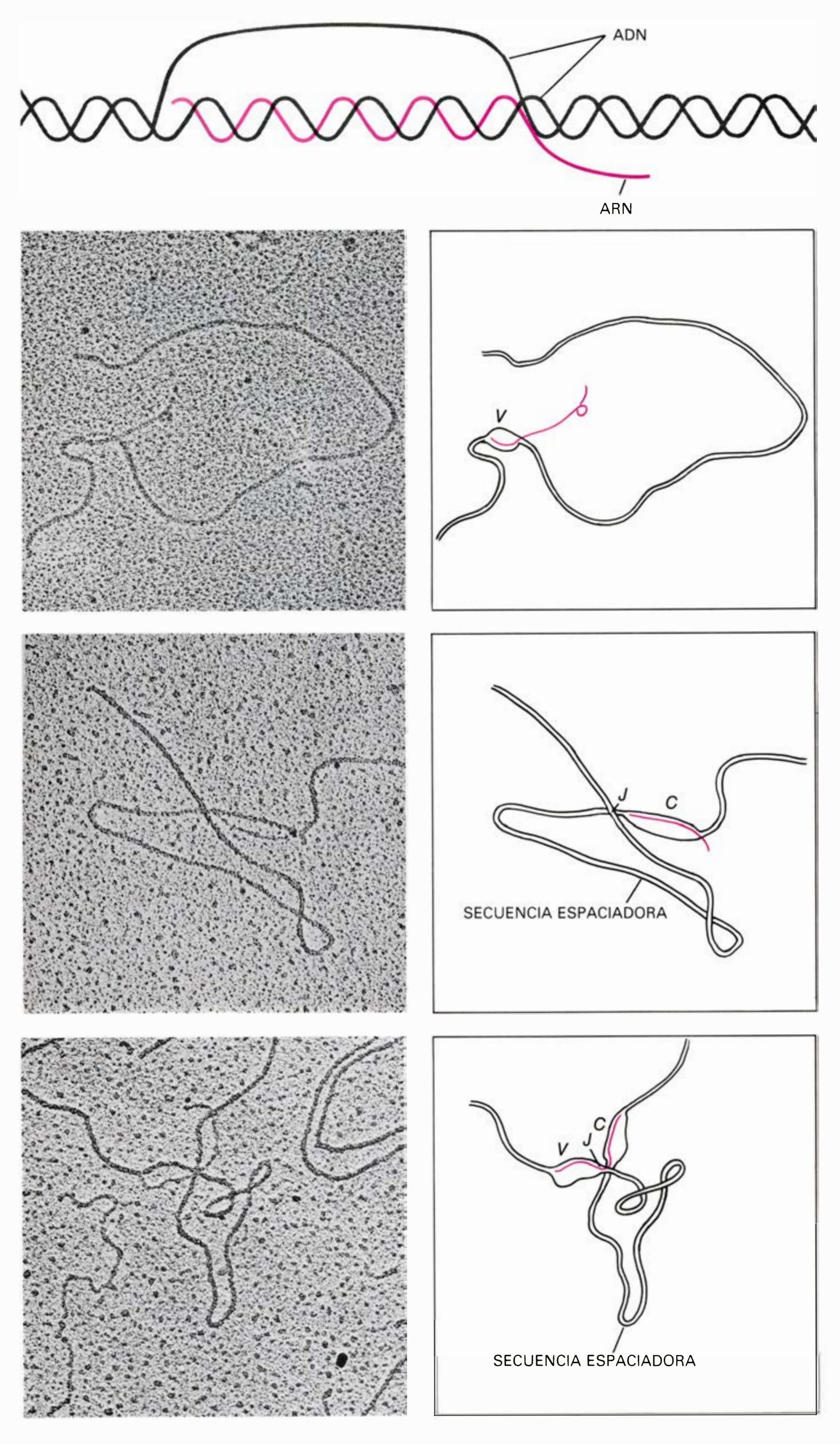
A mediados de los setenta se produjo un extraordinario desarrollo de la genética molecular, salpicado de controversias y malentendidos. Un avance fundamental, en 1973, fue la primera aplicación con éxito de las nuevas técnicas de ADN recombinante para insertar ADN extraño en una bacteria o un virus bacteriano y, de este modo, clonar un único gen en grandes cantidades. Los que trabajábamos con técnicas genéticas más laboriosas vimos claro que la clonación nos permitiría aislar genes de anticuerpos y determinar su estructura de forma directa. Había ciertas dudas, sin embargo, sobre los posibles riesgos de los procedimientos; la seguridad de los métodos ha quedado asentada, pero en aquel tiempo el Instituto Nacional de la Salud, estadounidense, promulgó unas cautelosas orientaciones para los experimentos de ADN recombinante. Entre los requisitos había unos que afectaban a los organismos vectores: la bacteria o virus bacteriano en los cuales se clonaba el ADN extraño. Los vectores debían "mutilarse" genéticamente para reducir por un factor de 100 millones su posibilidad de supervivencia fuera del laboratorio.

Mis colegas David C. Tiemeier, Lynn Enquist, Nathan Sternberg, Robert A. Weisberg y yo decidimos utilizar a este propósito el bacteriófago (o virus bacteriano) llamado fago lambda. Una versión del fago, preparada por Ronald W. Davis, de la Facultad de Medicina de la Universidad de Stanford, se manipuló genéticamente de forma que sólo creciese si incorporaba un segmento de ADN extraño. Introdujimos una serie de mutaciones "mutiladoras" en el fago vector, realizamos las pruebas apropiadas de supervivencia y, finalmente, recibimos el permiso pertinente. Muy pronto, Shirley M. Tilghman, Tiemeier y el autor informaron de la primera clonación con éxito de un gen de mamífero en un bacteriófago: un gen de globina de ratón (la proteína de la molécula de hemoglobina).

A solicitud del laboratorio de Tonegawa, enviamos a Suiza una copia de nuestro fago, donde Tonegawa y sus colaboradores aplicaban la estrategia de clonación para aislar los genes de ratón para la cadena ligera tipo lambda, tanto en células embrionarias como en las productoras de anticuerpos. Descubrieron que el gen de la región constante y los genes de la región variable estaban muy separados en el ADN de las células que no sintetizaban anticuerpos. En un plasmacitoma que producía anticuerpos los dos genes estaban mucho más próximos. La reorganización, sin embargo, no determinaba el empalme de los genes de las regiones variable y constante para formar una secuencia continua. Por el contrario, entre el gen de la región variable (*V*) y el gen de la región constante (*C*) mediaban todavía unos 1500 nucleótidos; había entre ellos, y colindando con el gen *V*, un segmento llamado secuencia *J*, del que nos ocuparemos más adelante.

La cadena kappa

Nuestros propios proyectos de clonación se encaminaron hacia los genes de la cadena ligera kappa de ratón. Nos movía a ello el hecho de que más del 95 por ciento de los anticuerpos de ratón incorporan la cadena kappa, constituyendo ésta la fuente más importante de diversidad de cadenas ligera de ratón. Jonathan G. Seidman y Edward E. Max, colegas míos, clonaron las formas embrionaria y activa de los genes de las regiones constante y variable de tipo kappa. Después, utilizando las técnicas de secuenciación rápida de ADN desarrolladas por Alan Maxam y Walter Gilbert, de la Universidad de Harvard,



LOS COMPONENTES DE LOS GENES DE LOS ANTICUERPOS pueden visualizarse en electromicrografías de ADN y ARN de ratón preparadas mediante una técnica especial de hibridación. Bajo ciertas condiciones se logra que una secuencia de ARN complementaria con una de las cadenas de un ADN bicatenario se hibride con ella más rápidamente de lo que pueda hacerlo la otra cadena del ADN; esta última se ve desplazada y forma un lazo, o bucle (*arriba*). En estas micrografías, el ARN (en color en los esquemas) es ARN mensajero para la cadena ligera; se trata de una secuencia *V/J/C* contigua. En la micrografía superior, se mezcló el ARN con ADN de células embrionarias, o de la línea germinal. La secuencia *V* del ARN se ha hibridado con una cadena de una secuencia *V* de la línea germinal, formando un bucle; el resto del ARN queda colgando. En la micrografía del centro, un bucle similar pone de manifiesto al gen *C*, cuyo ADN está separado de la secuencia *J* por una secuencia espaciadora larga y sin hibridar. (Se ha unido una cadena de la secuencia *J* de la línea germinal a un trozo de ARN mensajero no visible en la micrografía.) Abajo, el ARN mensajero se hibridó con ADN de una célula productora de anticuerpo, donde las secuencias *V* y *J* se han unido. En este caso, el ARN mensajero se hibrida con los tres genes, formando un bucle *V/J* y otro *C*; de nuevo, el ADN espaciador no se hibrida con el ARN.

determinaron la secuencia de nucleótidos de los genes.

En el ADN embrionario se identificaron un gran número de genes *V*. Parecen agruparse en familias, cada una de las cuales comprende genes cuyas secuencias nucleotídicas guardan una estrecha relación. Nuestros propios estudios, y los de Robert P. Perry, del Instituto para la Investigación del Cáncer en Fox Chase, Pennsylvania, indican que en el ADN embrionario de ratón puede haber varios centenares de genes de la región variable. Los genes *V*, tanto kappa como lambda, conservan ciertas características estructurales que parecen ser de considerable significación. Así, cada gen está dividido en dos segmentos informativos discretos, separados por una corta secuencia intercalada. La primera de las secuencias informativas determina la síntesis de un "líder" hidrofóbico (que repele el agua) de 17 a 20 aminoácidos de largo, que probablemente cumpla un papel importante en el transporte de la molécula de anticuerpo a través de la membrana celular. El líder es una parte del producto proteínico original del gen activo de la cadena ligera, pero se elimina cuando el anticuerpo primordial atraviesa la membrana.

La otra región informativa del gen *V* determina la síntesis de la mayor parte de la región variable, aunque no de toda. El segmento del gen *V* que no corresponde al líder determina sólo 95 de los 108 aminoácidos de la región variable de la cadena kappa. Como habían sugerido los primeros descubrimientos de Tonegawa para el caso del sistema lambda, encontramos que, en el ADN de kappa, la porción restante de la región variable está determinada por una secuencia que se halla "curso abajo" del gen *V*, cerca del gen de la región constante, exactamente en el sitio donde se unirá el gen *V* para formar un gen de inmunoglobulina activo. Esta corta secuencia, el gen *J*, se repite cinco veces, con ligeras y significativas variaciones, a intervalos de 300 nucleótidos, aproximadamente. (Estudios posteriores realizados en el sistema lambda del ratón por Tonegawa, David Baltimore, del Instituto de Tecnología de Massachusetts, y Ursula Storb, de la Universidad de Washington, han demostrado aquí una reorganización algo diferente. En vez de un gen para la región constante, hay cuatro genes *C*, cada uno de los cuales tiene su propio gen *J*. En el sistema lambda humano, Philip A. Hieter, Gregory F. Hollis y yo hemos encontrado seis genes *C*.)

Empieza pues a verse claro el poten-

cial del sistema kappa para generar diversidad. La unión de uno de los varios cientos de genes *V* –digamos 150 tirando por lo bajo– a uno de los cinco genes *J* puede constituir hasta 750 (150 × 5) genes activos diferentes para la región variable de la cadena ligera. Las pruebas acumuladas por numerosos laboratorios indican que esta es exactamente la forma en que se combinan las secuencias. Uno de los genes *V* se une a uno de los *J*; los demás genes *V* y *J* (y la larga secuencia intercalada no informativa), interpuestos entre ellos, se delecionan. El gen activo ya "acabado" consta de tres secuencias informativas separadas: un gen líder, un gen *V/J* y un gen *C*. Las secuencias se ensamblan durante la "maduración" del ARN y forman finalmente un ARN mensajero para cadena ligera correcto.

Esta diversidad, que se ha multiplicado ya por 750, se verá incrementada por otra fuente de variación. Una atenta comparación entre las secuencias de aminoácidos de las cadenas ligeras revela un grado de diversidad particularmente alto en una región cercana al sitio de unión *J/V*. Los aminoácidos situados alrededor de la posición 96 forman una de esas tres regiones que presentan las cadenas ligeras y que Elvin A. Kabat, de la Facultad de Medicina y Cirugía de la Universidad de Columbia, había denominado tiempo atrás "hipervariables". Las cadenas ligeras se pliegan de tal suerte que las regiones hipervariables constituyen los sitios de unión antígeno-anticuerpo.

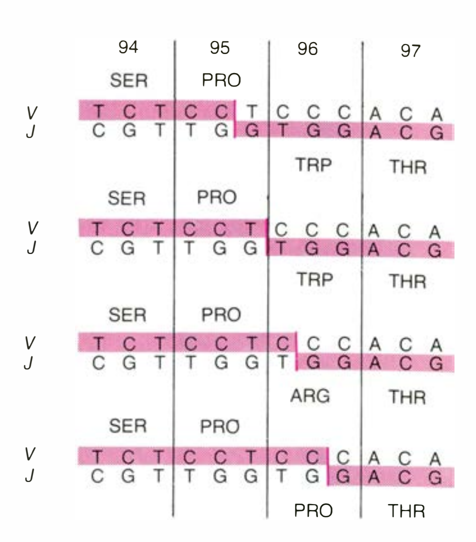
Parte de la variación en esa región puede explicarse ahora por el hecho de que el sitio de recombinación *V/J* no está definido con precisión. Aparentemente, un gen *V* y uno *J* pueden unirse de diferentes maneras [*véase la ilustración de esta misma página*], de modo que el codón para el aminoácido 96 (el punto de unión *V/J*) y los codones adyacentes pueden cambiar dependiendo de qué parte de la secuencia aporte la región *V* embrionaria y cuál la *J*. Si se acepta que los sitios de unión alternativos pueden incrementar unas diez veces la diversidad, el número total de combinaciones *V/J* potenciales llega a ser de 7500 (150 × 5 × 10).

Secuencias indicadoras

Mencioné antes que ciertas características de las secuencias de los genes de las cadenas ligeras se han conservado y que probablemente tengan una significación funcional. En particular, en los extremos 3' ("curso abajo") de los genes *V* y en los extremos 5' ("curso arriba") de los genes *J*, se han encontrado una serie de secuencias "indicadoras". Cada secuencia de éstas contiene un segmento de aproximadamente nueve nucleótidos (un nonámero) rico en adeninas (A) o timinas (T). A continuación del nonámero, y a intervalos de unos 11 o unos 22 nucleótidos, hallamos otra serie de siete nucleótidos, o heptámero: CACTGTG o GTGACAC. El nonámero y el heptámero pueden representarse en forma de una estructura en "tallo", donde las secuencias se complementan de acuerdo con las reglas de apareamiento (A se aparea con T y G con C), acercando los genes *V* y *J* a la base del tallo [*véase la ilustración superior de la página siguiente*]. Es fácil admitir, a partir de aquí, que los genes se unen mediante algún tipo de mecanismo de recombinación de ADN, eliminándose finalmente las secuencias indicadoras. En el proceso interviene, a buen seguro, un sistema específico (pero todavía desconocido) de enzimas de recombinación.

La flexibilidad del sistema de recombinación, aunque poderoso en su capacidad para generar diversidad, también paga su tributo. Ocasionalmente, los genes *V* y *J* se unen de forma incorrecta, dando lugar a un gen inactivo. Este hecho puede explicar en parte el fenómeno denominado exclusión alélica. Toda célula somática tiene dos series cromosómicas, en las que un miembro de cada par de cromosomas homólogos es aportado por la madre y el otro por el padre. Las correspondientes copias de un determinado gen en los dos cromosomas se llaman alelos. En una célula productora de anticuerpos normalmente ocurre que la recombinación somática afecta sólo a los genes de anticuerpos presentes en una de las copias del cromosoma que lleva dichos genes, y sólo ésos se traducen en proteínas; los alelos presentes en el otro cromosoma suelen quedar excluidos del proceso de reorganización y no se expresan. Pero ocurre a veces que una célula productora de anticuerpos posee dos series de genes de anticuerpos reordenados. Mis colegas y yo sospechamos que, en tales casos, los genes de uno de los cromosomas recombinaron erróneamente formando un gen inactivo y que, en un segundo intento, utilizando como apoyo el cromosoma de "recambio", se generó el gen activo.

Quizá los genes de cadena ligera lambda constituyan un sistema de seguridad para los genes kappa mal ensamblados. Estudios realizados por Hieter, Hollis, Stanley J. Korsmeyer, Thomas A. Waldmann y el autor revelan la existencia de una clara ordenación de acontecimientos durante la formación de un gen de cadena ligera. Primero se reorganizan los genes kappa. Si éstos no consiguen formar un gen activo, empieza a reorganizarse el sistema lambda.

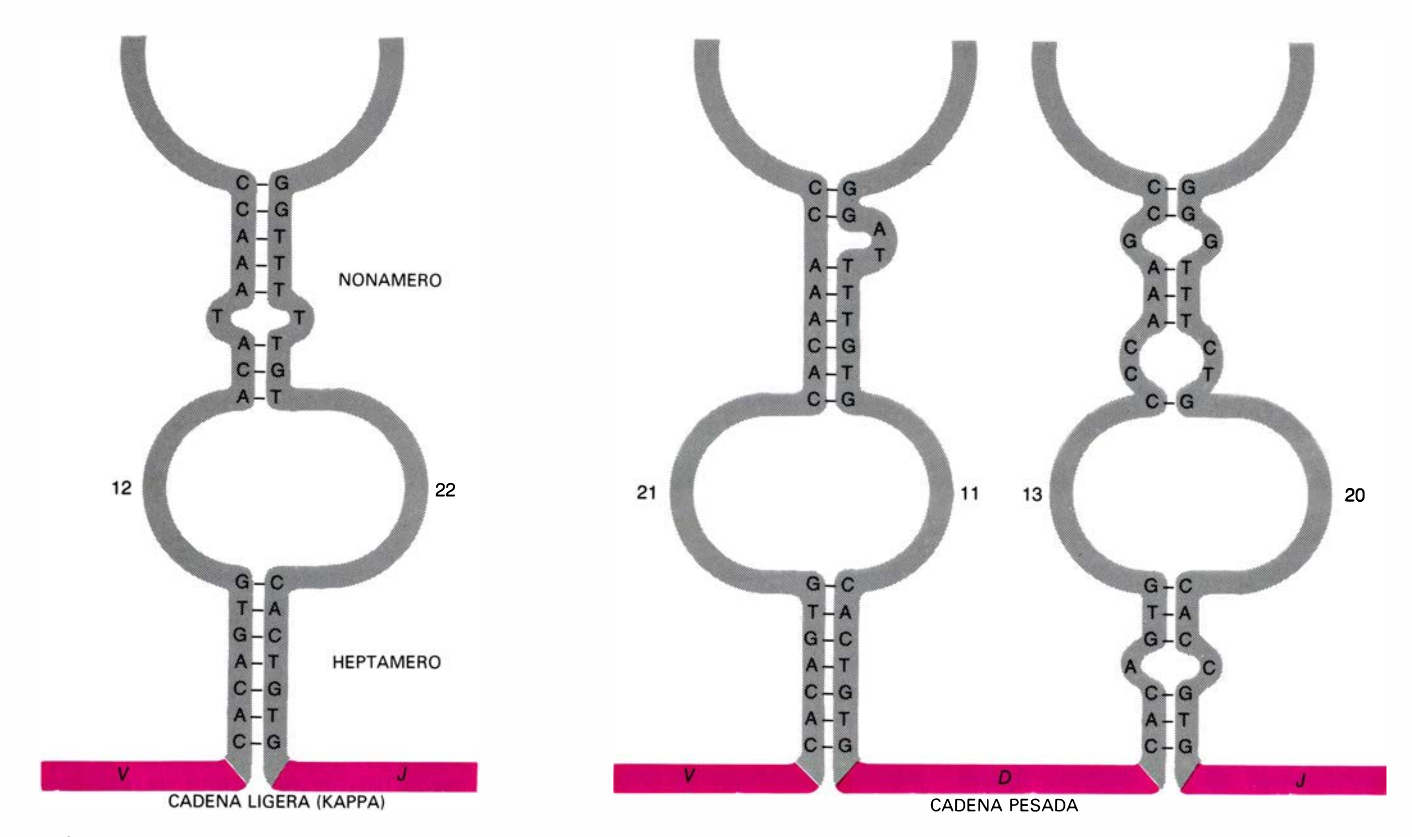


FLEXIBILIDAD DE LA RECOMBINACION, que amplía la diversidad de los anticuerpos. El punto de entrecruzamiento (*líneas de color oscuro*) en la recombinación entre las secuencias *V* y *J* puede variar en un margen de varios nucleótidos, dando lugar a diferentes secuencias nucleotídicas (*franjas de color*) en el gen activo de la cadena kappa. Resulta entonces que son varios los posibles codones para el aminoácido 96: TGG determina el aminoácido triptófano, CGG arginina y CCG prolina. La variación reside en la tercera región hipervariable de la cadena kappa, y puede ejercer un importante efecto en el sitio de unión al antígeno.

Baltimore y sus colaboradores, en el Instituto de Tecnología de Massachusetts, han propuesto la sugestiva idea de que la aparición de un anticuerpo funcional en una célula es como una señal que imposibilita posteriores uniones *V/J*; en ausencia de la señal, la célula sigue intentándolo. En los seres humanos, con seis secuencias *J/C* tipo lambda en cada cromosoma número 22, hay en realidad 12 genes lambda por célula. Junto con los dos genes kappa (uno en cada cromosoma número 2), la célula tiene 14 oportunidades para formar un gen de cadena ligera activo.

La cadena pesada

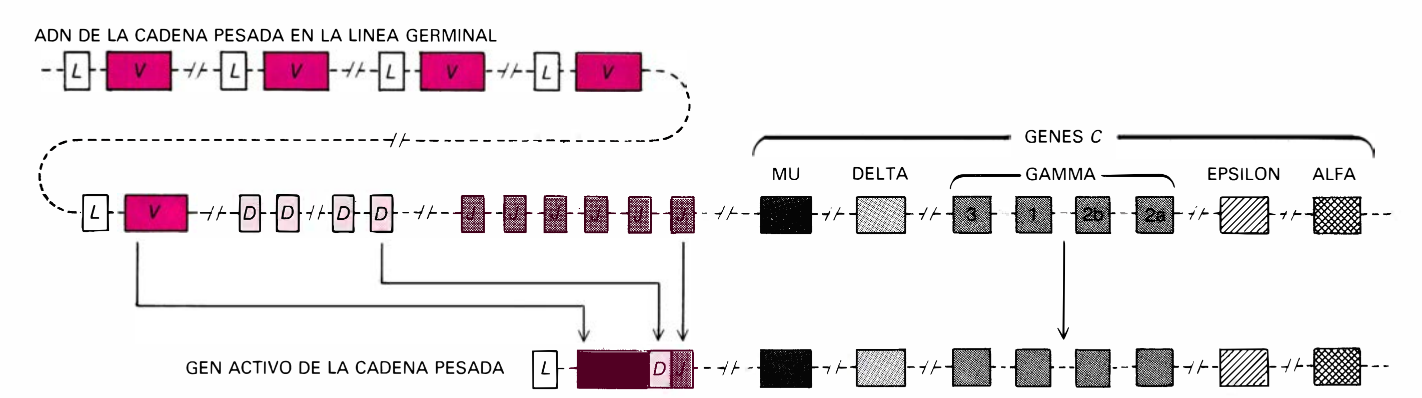
En la cadena pesada, la formación de la región variable sigue los mismos principios que se aplican en la cadena ligera, si bien el potencial de diversidad resulta todavía mayor: un segmento extra de información genética multiplica las posibilidades de combinación. Leroy E. Hood y sus colegas, del Instituto de Tecnología de California, clonaron ADN de la región variable de la cadena pesada activa de células productoras de



SEÑALES PARA LA RECOMBINACION en el ADN de la línea germinal; destínanse, al parecer, a acercar las secuencias *V* y *J* de la cadena ligera como paso previo a la recombinación. En el caso de la cadena pesada opera un mecanismo similar, que incluye además señales para una secuencia adicional llamada *D* (de diversidad). Cada señal consta de una secuencia de siete nucleótidos (casi siempre con una A o una T en el centro) y una secuencia de unos nueve nucleótidos rica en A o T. Entre heptámero y nonámero se intercala un espaciador de aproximadamente 11 o 22 nucleótidos de largo. Dos de esas señales, cuyos heptámeros y nonámeros se complementan extensamente según las reglas de apareamiento de bases (A se aparea con T y G con C), pueden formar una estructura a modo de tallo que permite el acercamiento de las secuencias informativas para que pueda haber recombinación. Por lo que se aprecia, esa recombinación somática sólo ocurre cuando una de las señales complementarias tiene un espaciador corto y la otra un espaciador largo.

anticuerpos. Al determinar su estructura, observaron una secuencia de al menos 13 nucleótidos situada exactamente en la unión *V/J* que no tenía contrapartida en los genes *V* o *J* del ADN embrionario. Pensaron que esta secuencia debía provenir de algún segmento del ADN embrionario, al que llamaron gen *D* (de diversidad). Vieron, además, que, en el gen activo, el segmento *D* correspondía a la parte principal de la tercera región hipervariable de la cadena pesada.

Philip W. Early y Hood descubrieron también en el ADN embrionario de la cadena pesada las secuencias nonámero-heptámero que en el ADN de la cadena ligera parecen servir de señales para llevar a cabo las uniones *V/J*. La disposición de aquellas señales en el ADN de la cadena ligera indicaba que, para que tuviera lugar la recombinación, la espaciación entre el nonámero y el heptámero debía ser diferente en los lados *V* y *J* de la estructura en tallo (alrededor de 11 nucleótidos en un lado y 22 en el otro). En la cadena pesada, sin embargo, se halló que las señales de *V* y las de *J* incluían espaciadores de alrededor de 22 nucleótidos. Early y Hood predijeron que, cuando se lograse identificar los genes *D* en el ADN embrionario, estarían flanqueados por



EL GEN ACTIVO DE LA CADENA PESADA surge de cuatro series de secuencias de la línea germinal (*arriba*): *L/V*, *D*, *J* y *C*. Mediante recombinación somática se empalman una de las secuencias *L/V*, una *D* y una *J*, formando la región variable. Como en el caso del gen de la cadena ligera, el ADN de la región constante (*C*) está situado curso abajo, separado por una secuencia no informativa. En la cadena pesada, sin embargo, hay ocho secuencias *C* separadas, cada una de las cuales determina la síntesis de una región constante diferente. (Cada secuencia *C* está dividida a su vez en otras cinco regiones, pero en el esquema se dibujan de forma simplificada.) El montaje final de las secuencias informativas se realiza durante el procesamiento del ARN.

señales de recombinación separadas por 11 nucleótidos. De esta forma, la unión de *V* a *D* y de *D* a *J* cumpliría la regla 11/22.

Los supuestos segmentos *D* presentaban un problema en la clonación. Su secuencia informativa era demasiado corta para utilizarla como sonda de detección, por hibridación, de su contrapartida embrionaria. Mis colegas Ulrich Siebenlist y Jeffrey V. Ravetch resolvieron el problema recurriendo a un producto intermedio anormal del proceso de recombinación *V*/*D*/*J*: un segmento *D* humano que se había unido incorrectamente a un gen *J*. Los enzimas de recombinación habían procesado sólo uno de los extremos del gen *D* (el adyacente a *J*); el extremo opuesto presentaba, por tanto, un gran segmento de ADN embrionario. Este segmento anormal constituyó una sonda excelente en la clonación de secuencias *D* embrionarias.

Cuando Siebenlist y Ravetch determinaron la estructura del ADN en esos clones, vieron que las secuencias *D* estaban rodeadas por señales de recombinación separadas por 11 nucleótidos, coincidiendo con las predicciones de la regla 11/22. A continuación, utilizando como sonda el segmento *D* clonado, Siebenlist y Ravetch rastrearon el genoma humano buscando secuencias complementarias. Detectaron una extensa familia de genes *D* formada por, al menos, cinco secuencias estrechamente relacionadas. Hay indicios de que ésta es sólo una de las varias familias de genes *D* que existen.

18.000 millones de anticuerpos

La formación del gen activo de la región variable de la cadena pesada puede generar un número extraordinariamente grande de posibilidades genéticas. T. H. Rabbitts y sus colegas, del Laboratorio de Biología Molecular del Consejo de Investigaciones Médicas de Cambridge, han estimado que en el hombre hay al menos 80 genes *V* embrionarios de cadena pesada. Ravetch ha encontrado seis genes *J* activos en 8000 nucleótidos del gen *C* mu humano. Aunque no se pueda extrapolar con seguridad el número de genes *D* humanos a partir de lo conocido hasta ahora, asumiré que las familias *D* se componen de unos 50 miembros. Así, la recombinación somática puede generar, aproximadamente, unas 24.000 combinaciones genéticas (80 × 6 × 50). El cálculo puede multiplicarse por un factor de 100 aproximadamente (una estimación poco fina) debido a la flexibilidad recombinatoria: codones alternativos en los dos puntos de entrecruzamiento de los procesos de recombinación *V*/*D* y *D*/*J*. El total es, por tanto, de aproximadamente 2,4 millones de posibles cadenas pesadas diferentes.

Multiplicando las 7500 posibilidades combinatorias para el caso de la cadena ligera kappa humana (150 por el complemento de genes *V*, 5 por los genes *J* y 10 por flexibilidd recombinatoria) por los 2,4 millones de cadenas pesadas, resulta un total de 18.000 millones de an-

DIFERENCIACION DE LAS CELULAS productoras de anticuerpos a partir del linfocito pre-*B*. En cada estadio se especifican los sucesos genéticos más importantes (*izquierda*), junto con los correspondientes mecanismos de mezcla génica (*derecha*). Primero se fabrica la cadena pesada y después la cadena ligera (kappa o lambda). Las moléculas de IgM o IgD se disponen en la superficie celular. Cuando uno de esos anticuerpos de superficie reconoce y se une a un antígeno específico, se estimula el desarrollo de la célula: prolifera hasta formar un clon de linfocitos *B* maduros, especializados en la síntesis de grandes cantidades de proteína. Durante el proceso puede cambiar la expresión y ordenación de los genes de la región constante de la cadena pesada, fabricándose diferentes tipos de anticuerpos. En la mayoría de los casos las regiones variables son las mismas y el anticuerpo actúa contra el mismo antígeno, pero durante el proceso de maduración pueden acumularse mutaciones puntuales que alteren las regiones variables.

ticuerpos posibles (2,4 millones × 7500), generados a partir de unos 300 segmentos genéticos distintos, localizados en el ADN embrionario.

Esa enorme diversidad generada mediante recombinación se amplía si consideramos otro mecanismo: mutaciones somáticas aisladas, que introducen cambios esporádicos de nucleótidos a lo largo del ADN de la región variable durante el desarrollo somático. Los genes de las inmunoglobulinas son altamente inestables en las células productoras de anticuerpos. Matthew D. Scharff y sus colegas, de la Facultad de Medicina Albert Einstein, al cultivar clones de linfocitos maduros y analizar sucesivas generaciones para nuevas especificidades antigénicas, descubrieron en los genes de las inmunoglobulinas una tasa de mutación considerable: una por cada 10.000 células y por generación. En otros experimentos se aislaron genes activos en los que el ADN de la región variable difería en uno o dos nucleótidos de las secuencias de ácido desoxirribonucleico embrionario de las que procedían.

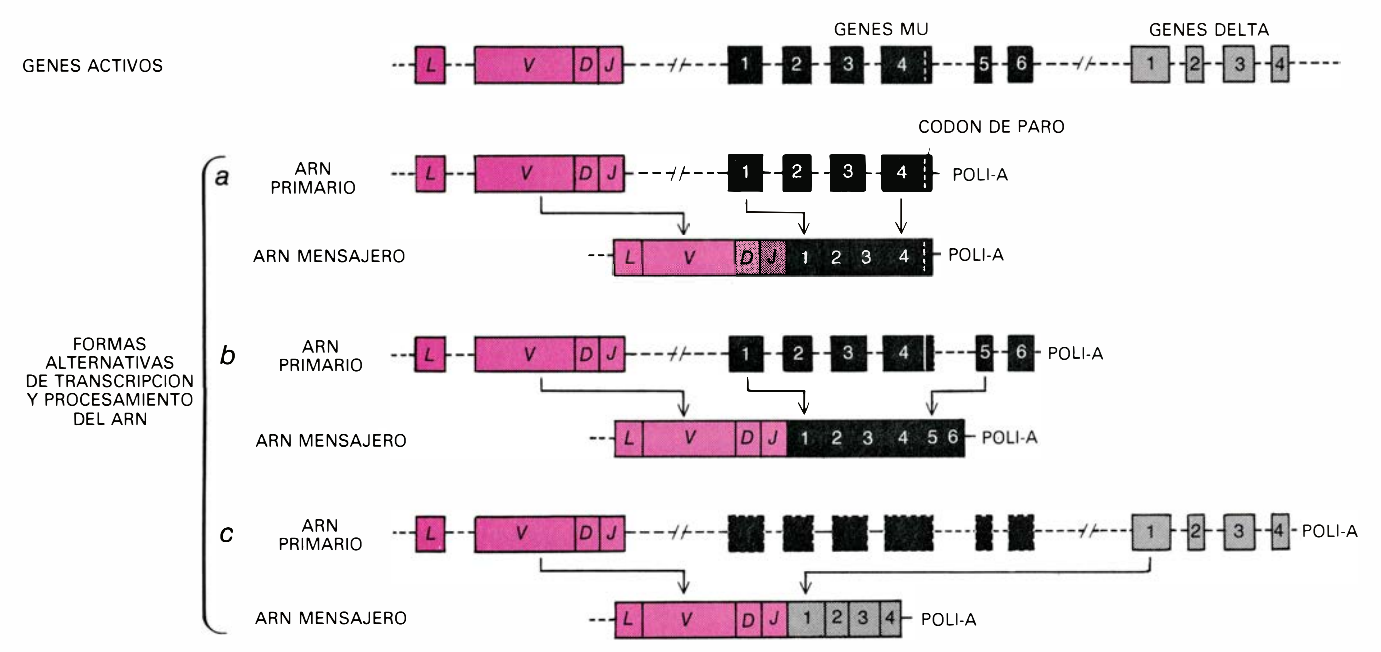
Estos hallazgos apoyan la propuesta avanzada hace unos 15 años por Melvin Cohn y Martin Weigert, del Instituto Salk de Estudios Biológicos, de que la diversidad de los anticuerpos resulta en parte de mutaciones de cambios de bases en el ADN de la región variable durante el desarrollo de los linfocitos. Patricia J. Gearhart, de la Institución Carnegie de Washington, y Hood y Baltimore y sus colaboradores tienen pruebas suficientes para sugerir que las mutaciones se acumulan durante los diferentes estadios del desarrollo de los linfocitos. Se desconoce el mecanismo que provoca tales mutaciones.

La mezcla de genes capaz de generar miles de millones de genes para la región variable va acompañada de dos procesos adicionales que explican de qué manera una misma región variable puede unirse sucesivamente a una serie de regiones constantes de la cadena pesada. El linfocito pre-*B*, precursor de las células productoras de anticuerpos, fabrica una región constante de cadena pesada tipo mu unida a una región variable específica (producto de la recombinación *V*/*D*/*J*). Esta cadena pesada permanece en principio dentro de la célula pre-*B*. Más tarde, tras la puesta en marcha de la síntesis de cadenas ligeras y cadenas pesadas delta, las pesadas delta y mu se combinan con las cadenas ligeras para formar moléculas completas de IgM e IgD. El siguiente estadio del desarrollo se caracteriza por la aparición conjunta en la superficie celular de moléculas de IgM e IgD. Ambos anticuerpos tienen las mismas regiones variables y actúan, por tanto, contra el mismo antígeno.

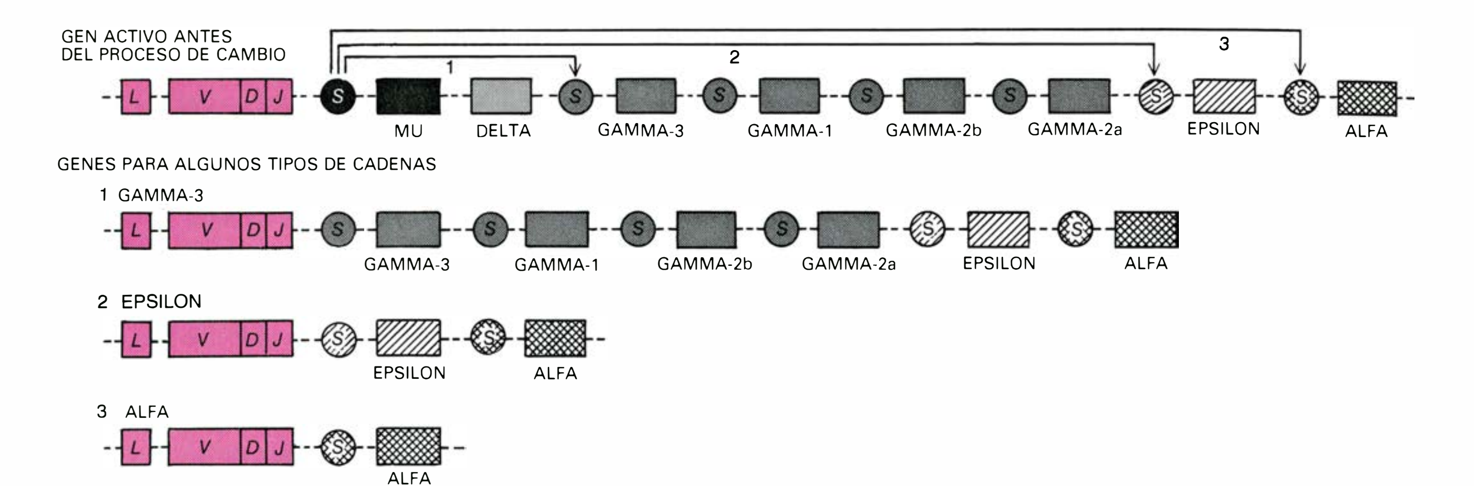
Las siguientes etapas de la maduración de los linfocitos están dirigidas, según parece, por el antígeno. Lo primero que ocurre en el proceso inmunológico es la selección clonal. Un antígeno se une a un receptor: el lugar de unión que mejor se ajusta entre millones o miles de millones de inmunoglobulinas de superficie. La interacción estimula el desarrollo de la célula que posee la inmunoglobulina seleccionada, que prolifera y forma un clon de linfocitos *B* productores de anticuerpos.

Selección clonal

Durante la maduración de las células *B*, las moléculas de IgM e IgD desaparecen de la superficie celular y la célula excreta IgM, IgG, IgE o IgA. Aunque cada una de esas inmunoglobulinas posee una cadena pesada con una región constante diferente, los anticuerpos de cualquier clase que se sintetizan en una célula evidencian las mismas regiones variables, aquellas que se formaron en la célula precursora y que constituyen el sitio de unión seleccionado por el antígeno. Puesto que cada cadena pesada confiere al anticuerpo una función efectora peculiar, el mismo sitio de unión puede participar en reacciones



LA TRANSCRIPCION Y EL PROCESAMIENTO DEL ARN explican la aparición sucesiva de moléculas de IgM unidas a membrana y excretadas, así como la aparición simultánea de IgM e IgD. La línea superior presenta la configuración de los genes de la cadena pesada, incluyendo la región variable y las regiones constantes mu y delta, en una célula productora de anticuerpos. Si la transcripción acaba en la cuarta región informativa del gen mu (*a*), el ARN, una vez procesado, da lugar a un ARN mensajero específico de una cadena a la que le falta una breve secuencia de aminoácidos, que es la que permite a la cadena anclarse en la pared celular; la molécula de IgM fabricada a partir de este ARN mensajero se excreta. Si el ARN primario comprende las regiones 5 y 6 de mu (*b*), durante el procesamiento que origina el ARN mensajero se elimina el codón de paro que se encuentra al final de la región 4, y se añaden las dos últimas regiones, que son las que determinan la síntesis de la secuencia de anclaje; estas moléculas de IgM permanecen entonces unidas a la membrana. Si el ARN primario incluye el gen delta (*c*), el procesamiento elimina a veces el gen mu y empalma las regiones delta directamente a *J*, en la secuencia de la región variable, fabricándose un ARN mensajero que determina una cadena delta y, por tanto, IgD. (Los dibujos son de Bunji Tagawa.)



CAMBIO DE CLASE DE CADENA PESADA, que se realiza mediante recombinación del ADN. Mientras un linfocito sintetiza cadenas mu o delta, el gen de la cadena pesada presenta la ordenación especificada en la parte superior. (En realidad, las secuencias que determinan cada región constante están divididas en varias regiones separadas por largos segmentos no informativos.) A cada gen de región constante precede una señal de cambio (*S*), que guarda alguna relación de complementariedad con otra señal similar presente entre las secuencias de región variable y el gen mu. Parece que estas señales intervienen en un proceso de recombinación que une la secuencia *V/D/J* a una de las secuencias de la región constante situada curso abajo; la figura ofrece tres alternativas. Decidida la clase, el ADN se transcribe y el ARN se procesa, formándose un mensajero específico para la región gamma-3, épsilon o alfa.

inmunológicas distintas. El proceso mediante el cual la misma región variable aparece asociada con diferentes regiones constantes podría denominarse "cambio de clase de cadena pesada". El linfocito cuenta con dos mecanismos para llevar a cabo ese cambio de clase; uno basado en la transcripción diferencial del ADN y posterior procesamiento del ARN y, el otro, en un proceso de recombinación del ADN, fundamentalmente.

En los tres últimos años, varios grupos de investigadores han conseguido descifrar la organización, en una extensión de más de 100.000 nucleótidos, de los genes de las cadenas pesadas del ADN embrionario de ratón. Hood y el grupo de Tonegawa, con éste al frente, han contribuido a dilucidar la estructura de los genes mu, gamma y alfa. Frederick R. Blattner y sus colegas, de la Universidad de Wisconsin en Madison, han establecido la localización del gen delta. Tasuku Honjo y sus colegas, en la Facultad de Medicina de la Universidad de Osaka, han clonado y cartografiado los genes gamma (hay cuatro subclases de IgG), empalmándolos entre sí y a los genes épsilon y alfa. La disposición de estos genes (leyendo en el sentido de la transcripción, esto es, de 5' a 3') es: mu, delta, gamma 3, gamma 1, gamma 2*b*, gamma 2*a*, épsilon y alfa.

Forma doble de IgM

¿Por qué las IgM pueden aparecer, sucesivamente, de dos formas: unidas a la membrana del linfocito o excretadas? Jonathan W. Uhr, del Centro de Ciencias de la Salud de la Universidad de Texas en Dallas, observó una diferencia estructural entre las cadenas mu unidas a membrana y excretadas. Estudios más detallados han señalado que las formas unidas a membranas tienen en su extremo una corta secuencia de aminoácidos hidrofóbicos que sirven para anclar el anticuerpo en la membrana celular, debido a su afinidad por los lípidos hidrofóbicos de la membrana. La cadena mu excretada carece de esa secuencia hidrofóbica. Hood, con Randall Wall, de la Universidad de California en Los Angeles, y Baltimore y sus colegas, del MIT, han demostrado que se pueden sintetizar dos formas de ARN mensajero a partir del gen mu. En una de ellas, el mensaje se detiene muy cerca de los dos pequeños segmentos que determinan la secuencia de anclaje; en la otra forma, se incluyen los segmentos que codifican dicha secuencia.

Cada una de las regiones constantes de la cadena pesada se forma a partir de tres a seis regiones informativas distintas y separadas por breves secuencias espaciadoras de ADN no informativo. Como ocurre con la mayoría de los genes fragmentados, el ARN primario transcrito a partir de los genes de la cadena pesada incluye tanto las secuencias informativas como las espaciadoras. Durante el procesamiento del ARN se unen las secuencias informativas, eliminándose los segmentos espaciadores. El ARN primario del gen mu incluye a veces la secuencia que determina la parte hidrofóbica (el ancla); en presencia de ésta, los enzimas que procesan al ARN actúan sobre el extremo del mensajero de mu eliminando un codón de paro que, de otra forma, podría detener la traducción al final del mensaje principal. Si el ARN primario no posee la secuencia de anclaje, no hay tal procesamiento; el codón de paro que va delante de la secuencia de anclaje se reconoce como tal y se sintetiza la forma excretable de la cadena mu [*véase la ilustración de la página precedente*]. Hay pruebas cada vez más convincentes de que muchas de las demás cadenas pesadas, si no todas, muestran una ordenación similar, que les permite permanecer ancladas en la membrana o ser excretadas.

La aparición simultánea de cadenas delta y mu probablemente se deba a alternativas de procesamiento similares. La transcripción avanza por el ADN de la región variable y ataca luego las distintas regiones del gen mu. Cierta fracción del ARN primario acaba de forma que produce cadenas mu como las descritas anteriormente. Otra fracción del ARN primario continúa unos pocos miles de nucleótidos más, incluyendo no sólo las secuencias *V/D/J* y mu sino también la delta. Entre las muchas opciones existentes para procesar tales ARN, una consiste en unir el ADN *V/D/J* directamente con el comienzo de la región delta. Se forman así dos ARN mensajeros simultáneamente; uno de ellos con la información de la cadena pesada mu y otro con la de delta, pero ambas con la misma región variable. La traducción de los dos ARN en proteína lleva a la presencia conjunta en la superficie celular de IgM e IgD con la misma especificidad antigénica.

Un segundo mecanismo que permite combinar genes y llevar a cabo los cam-

bios de clases de cadena pesada se basa en la reordenación de secuencias de ADN. En contraste con la más que precisa naturaleza de los mecanismos de unión *V/J* y *V/D/J*, los procesos de recombinación que determinan la expresión de una clase particular de cadena pesada tienen un grado de libertad mucho mayor.

Consideremos el cambio de la cadena mu a la cadena alfa (de IgM a IgA). El gen activo de cadena pesada está organizado de tal manera que la secuencia que determina la región variable queda bastante alejada (aproximadamente unos 8000 nucleótidos) de la primera de las secuencias de la región constante (la primera secuencia mu). Normalmente, esta región espaciadora no informativa se elimina en el ARN primario de la cadena mu. Aunque la región espaciadora no posee una función informativa conocida, presenta un segmento de aproximadamente 2000 nucleótidos que, como Hood y Tonegawa y su grupo han demostrado, abarca una serie de secuencias de nucleótidos repetidas de varios tamaños.

En nuestro laboratorio, Ravetch e Ilan R. Kirsch han puesto de manifiesto que el segmento de 2000 nucleótidos con los bloques de secuencias repetidas es similar en secuencia a un segmento de ADN sito bastante más bajo, al lado del gen alfa. El cambio de IgM a IgA debe implicar recombinación entre esas secuencias similares distantes; por lo que se ve, sirven de señales para unir la secuencia *V/D/J* a la secuencia alfa, previa deleción de mu y los demás genes de la región constante. Ya que no hay secuencias informativas en estas regiones indicadoras, los puntos exactos de entrecruzamiento en la recombinación pueden en principio variar ampliamente dentro de esa zona. Tanto Tonegawa como Honjo han encontrado regiones indicadoras análogas al lado de los genes gamma. Sugieren que esas regiones representan una señal para el cambio de mu a gamma.

Consecuencias generales

No hace muchos años, los genéticos y biólogos moleculares aceptaban como axiomáticos los principios de que toda cadena proteínica estaba determinada por un gen particular de su línea germinal, que el genoma total de las células embrionarias permanecía inalterable mientras las células somáticas se diferenciaban hasta formar tejidos y que las formas y funciones características de las células diferenciadas dependían de la expresión diferencial de los mismos genes. El descubrimiento de que el sistema inmunológico genera muchos millones o miles de millones de proteínas combinando unos pocos cientos de genes de la línea germinal demuestra que esos principios no siempre se satisfacen. Quizá otras proteínas, además de los anticuerpos, requieran también más información que la que le puede ofrecer un genoma lineal inalterado, y parece probable, por tanto, que la recombinación somática del ADN y la transcripción selectiva y el procesamiento del ARN sean fuente de diversidad en otras células.

En un experimento realizado recientemente en mi laboratorio por Kirsch, Ravetch, Max y Robert L. Ney se hallaron indicios de que tales mecanismos podrían utilizarse fuera del sistema inmunológico. Emplearon el cambio de clase mu-alfa como sonda con la que buscar segmentos similares de ADN en genomas de ratón y humanos. Cada sonda reconoció de 10 a 15 fragmentos de ADN capaces de hibridarse con las señales de recombinación, lo que resultaba excesivo para las necesidades de un cambio de clase de cadena pesada. Se aplicó entonces la sonda a ADN de linfocitos en los que en ambos cromosomas se habían eliminado, por deleción, todos los genes de cadena pesada situados entre los genes *V* y el último gen de la región constante. Nuevamente, la sonda reconoció de 10 a 15 fragmentos, demostrando que las señales "de más" se encontraban fuera de las secuencias de las regiones constantes; es decir, en alguna parte del aparato genético del organismo.

Se sigue de ello que dichas señales de cambio de clase median la recombinación somática de otros genes, para otras proteínas que requieren un alto grado de diversidad. Por ejemplo, la recombinación podría servir para generar las proteínas que se manifiestan en la superficie de las células embrionarias y que parecen guiarlas hacia un lugar y destino específicos dentro de un organismo complejo; o quizá para generar la diversidad antigénica de la población de receptores localizados en la superficie de las células *T*, un tipo de linfocitos reguladores. Tales posibilidades están aún por estudiar. De lo que no cabe duda es de que el sistema inmunológico prueba la gran potencialidad de las redistribuciones de genes, que incrementan la información genética disponible para fabricar anticuerpos, y que probablemente sean piedra angular del desarrollo de diversidad génica.

La asimetría de las platijas

Los peces planos presentan ambos ojos en el mismo lado de la cabeza, predominando en la mayoría de las especies uno de los dos lados. ¿Qué fundamento adaptativo tiene esa preferencia "dextrógira" o "levógira"?

David Policansky

Casi todos los vertebrados poseen una simetría bilateral o próxima a ella: las estructuras a la izquierda y a la derecha del eje del cuerpo son imágenes especulares. Entre las excepciones más sorprendentes a esa regla se cuentan los peces planos, como la platija. En los peces planos adultos, se observan muchas desviaciones del esquema bilateral del cuerpo, siendo la más llamativa de ellas que ambos ojos están en el mismo lado de la cabeza.

La asimetría de los peces planos surge durante el proceso de desarrollo. Cuando el pez eclosiona del huevo su forma es simétrica, pero los rasgos y posición de las aletas y otras estructuras del cuerpo se modifican a medida que atraviesa las distintas etapas larvarias. Finalmente, un ojo migra por la parte superior de la cabeza y el pez adquiere el hábito de posarse en el fondo, con ambos ojos en la cara superior.

Además de esa asimetría, digamos individual, existe otra asimetría, igualmente sorprendente, a nivel de especie. En algunas especies, casi todos los peces adultos reposan sobre el flanco derecho y tienen ambos ojos en el lado izquierdo de la cabeza; se dice que son levógiros, o que muestran asimetría izquierda. En otras especies, prácticamente todos los adultos ofrecen la asimetría opuesta. Esta uniformidad manifiesta en la mayoría de las especies (aunque no en todas) sugiere que la asimetría puede estar bajo control genético y sujeta a la influencia de la selección natural. Si ello es así, cabe preguntarse: ¿qué ventaja selectiva supone para el pez posar sobre un flanco con preferencia al otro?

Según el criterio taxonómico que elijamos, el orden de los peces planos, los Pleuronectiformes, abarca seis o siete familias, que comprenden unas 500 especies. Como la mayor parte de los peces de interés comercial, los peces planos tienen nombres comunes que a veces pueden inducir a confusión. De entre las especies más conocidas citemos la platija de invierno, la platija de verano, el rodaballo, el lenguado inglés. el fletán, la platija estrellada y la solla. Ahora bien, mientras que en Inglaterra el lenguado se considera miembro de la familia de los lenguados, el pez llamado en California lenguado inglés pertenece a la familia de las platijas y, en toda Norteamérica, lo que la mayoría de los restaurantes ofrecen como filete de lenguado suele ser platija. Aquí emplearé el nombre de platija para referirme a los peces planos en general.

Desde el punto de vista de la biología, reviste especial interés la platija estrellada, *Platichthys stellatus*. Sus parientes más próximos son, en su mayor parte, dextrógiros; no así ella, que presenta tanto individuos levógiros como dextrógiros. La especie se encuentra en las aguas costeras del Pacífico, desde el sur de California hasta el Japón. En la costa oeste de los Estados Unidos las estrelladas son, aproximadamente, la mitad levógiras y la otra mitad dextrógiras. A medio camino entre los Estados Unidos y Japón, en las costas de Alaska, alrededor del 70 por ciento de ellas son levógiras. En las aguas japonesas casi el 100 por ciento son levógiras.

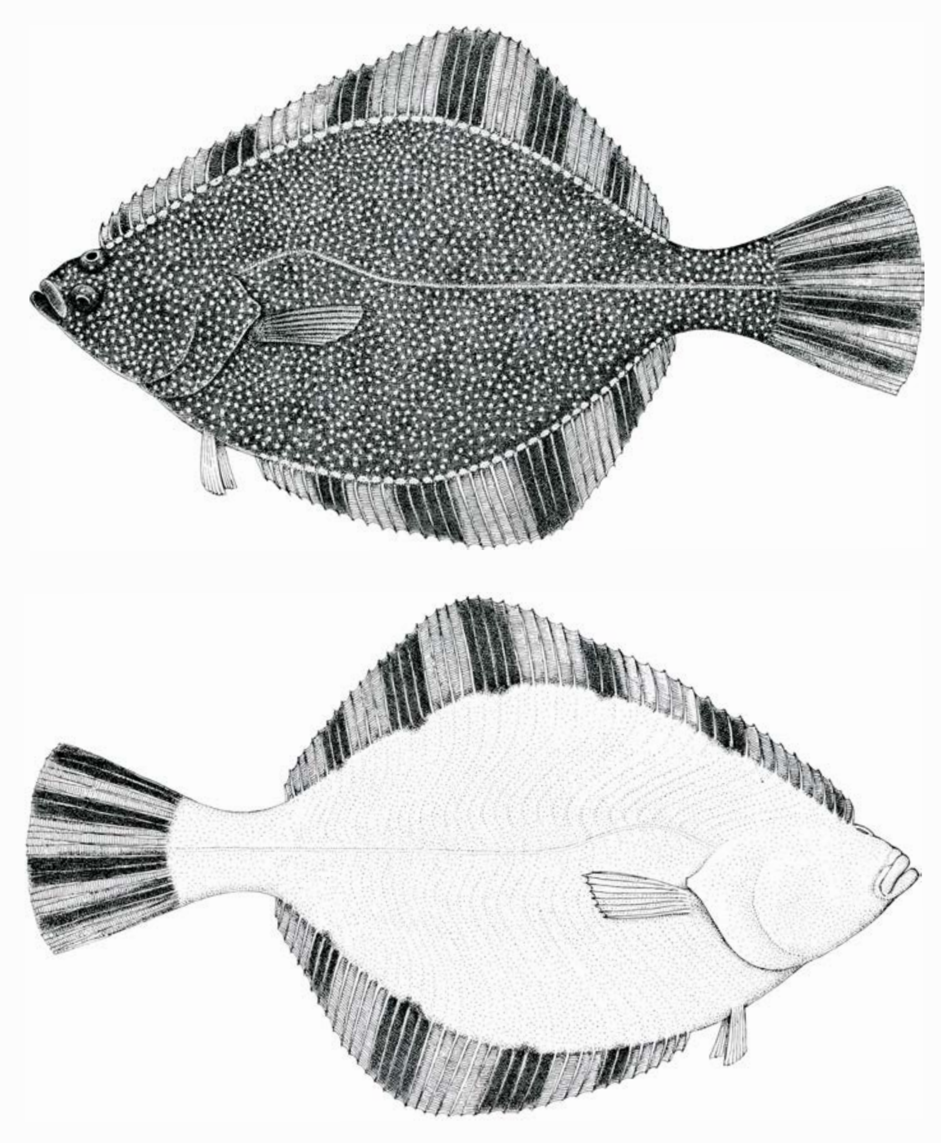
A este tipo de variación gradual sobre una cierta extensión geográfica, los biólogos lo denominan clina. La importancia de las clinas reside en su naturaleza indicativa de selección natural, o de diferencias de "adaptatividad". En el caso de la platija estrellada, la diferencia es la que hay entre imágenes especulares. Llama la atención que tal diferencia geométrica pueda tener valor adaptativo; de ahí que nos preguntemos si la clina constituye un accidente histórico o si interviene en ella la selección natural. Para plantearse siquiera una pregunta así debemos saber primero si el fenómeno está bajo control genético.

A pesar de lo que se sabe ya sobre el modo de acción de los genes, tanto en casos generales como en muchos ejemplos concretos, nadie puede explicar aún cómo se codifican las asimetrías en el material hereditario. Basta echar una ojeada a la confusa bibliografía sobre la dominancia de zurdos o diestros en el caso del hombre, para advertir cuán difícil resulta decidir siquiera que la direccionalidad de la asimetría esté bajo control genético. De ahí mi interés por estudiar la asimetría izquierda y derecha de la platija estrellada a la luz de fenómenos similares en otros organismos y comparar los resultados con lo que predicen dos modelos de herencia de las asimetrías. Analizaré primero los dos modelos y nos adentraremos luego en el estudio de las platijas estrelladas japonesas y americanas.

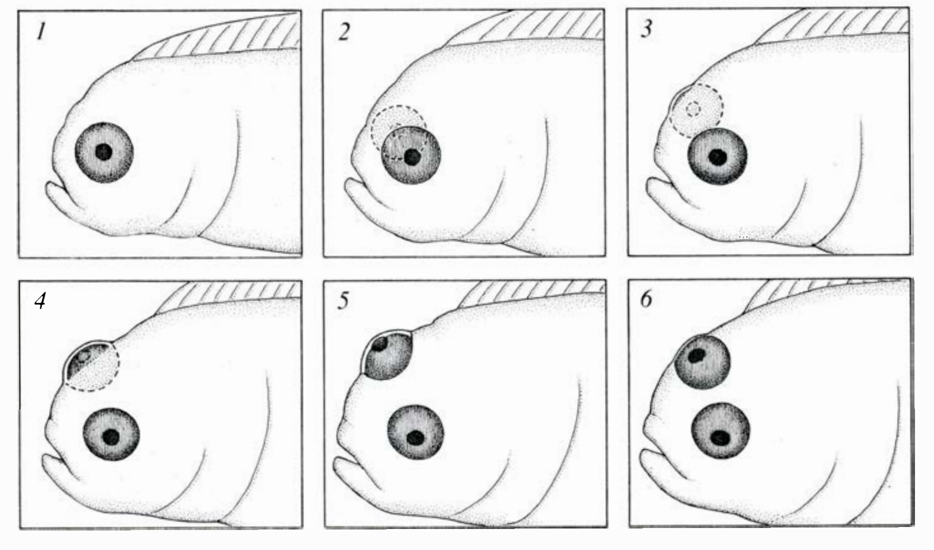
Los ejemplos más claros de control genético de asimetría izquierda y derecha los hallamos en tres especies de caracoles: *Lymnaea peregra*, cuya concha se enrolla normalmente hacia la derecha, *Laciniaria biplicata*, cuya concha se arrolla hacia la izquierda, y *Partula suturalis*, que puede tener la concha enrollada tanto hacia la derecha como ha-

LOS CUARENTA Y OCHO PECES pintados a la acuarela de la página opuesta son platijas estrelladas, *Platichthys stellatus*. Habitan en las aguas costeras del Pacífico, desde California hasta el Japón. Los aquí representados por el artista californiano Albert M. Plaza son miembros de la población de esa región estadounidense. Un sesenta por ciento de ellos tienen ambos ojos en el lado izquierdo de la cabeza; los restantes los tienen en el derecho. Menos del uno por ciento de las platijas estrelladas de aguas japonesas tiene ojos "dextrógiros". Los apareamientos experimentales efectuados entre las dos poblaciones han permitido evaluar diversas hipótesis sobre el control, genético o no, de esa curiosa asimetría ocular.





PLATIJA ESTRELLADA DE OJOS LEVOGIROS, según aparecería a un observador que la contemplara posada sobre el fondo (*arriba*) y vuelta del revés (*abajo*). Es una de las 500 especies de peces planos.



EMIGRACION DE LOS OJOS, que comienza cuando las larvas de las platijas estrelladas tienen unos 7 milímetros de talla; transforma en el pez joven su condición simétrica original (*1*). El ejemplar que se ilustra aquí está camino de ser de ojos levógiros. El ojo del lado distal (derecho) comienza a desplazarse rostro arriba hacia la parte superior del cráneo (*2, 3*). El pez es todavía transparente, y el cambio del ojo lejano se aprecia si se observa el animal desde el lado proximal (izquierdo). El proceso continúa (*4, 5*) durante cinco días, hasta que el ojo migratorio se ha desplazado casi completamente al lado proximal (*6*).

cia la izquierda. Otros ejemplos de asimetrías controladas genéticamente vienen dados por un insecto (una cepa mutante de laboratorio de la mosca del vinagre *Drosophila*), un protozoo (una rata mutante de *Tetrahymena*) y dos especies de plantas estrechamente relacionadas que pertenecen al género *Medicago*, grupo de leguminosas que incluye a la alfalfa común.

En las tres especies de caracoles, la constitución genética de la hembra determina el sentido de enrollamiento de la descendencia; el enrollamiento hacia la derecha es genéticamente dominante en *Lymnaea* y el enrollamiento hacia la izquierda lo es en las otras dos especies. En el mutante de *Drosophila* el abdomen se retuerce, bien en el sentido de las agujas del reloj, bien en sentido contrario; las diferentes direcciones del giro están determinadas por variantes génicas de un mismo locus de un cromosoma de *Drosophila*. Cada variante génica constituye un alelo distinto, de forma que en el locus "abdomen retorcido" del cromosoma puede haber un alelo "izquierdo", un alelo "derecho" o quizás un alelo "neutro".

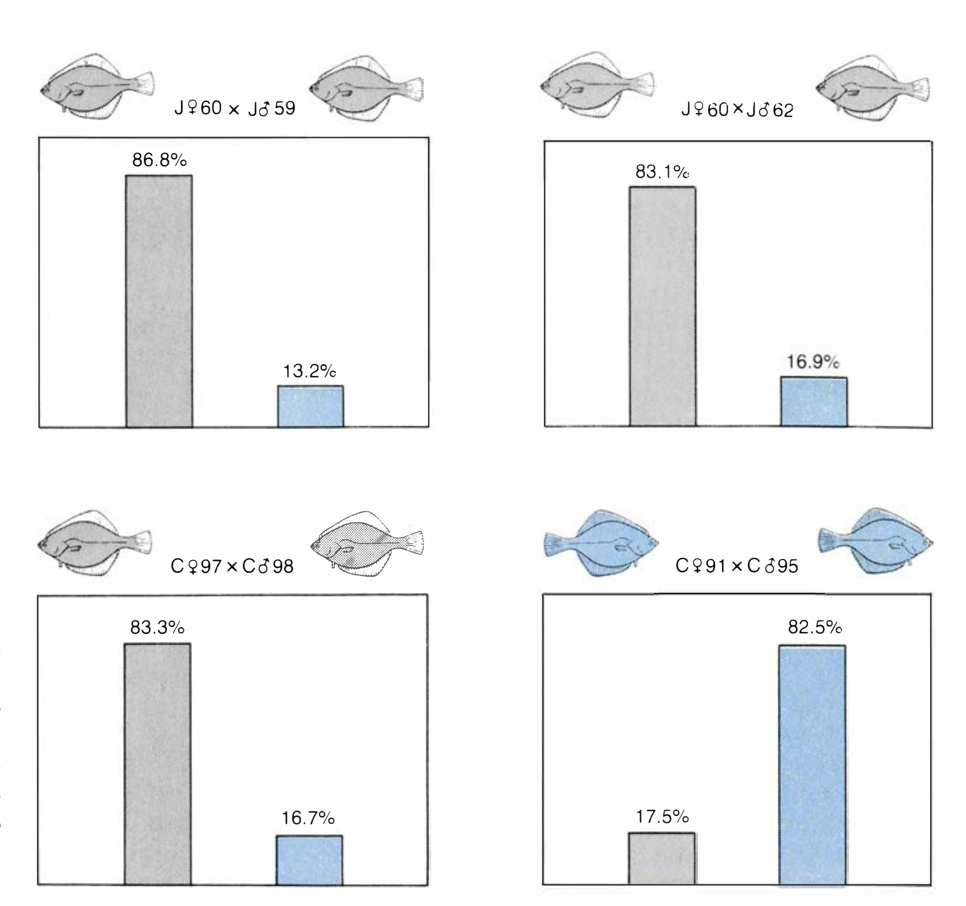
La mayoría de animales y plantas tienen los cromosomas a pares. Cada locus del cromosoma dispone de dos alelos, uno procedente del padre y el otro de la madre. Los alelos de un par pueden ser iguales, en cuyo caso el organismo es homicigótico; y pueden ser diferentes, en cuyo caso es heterocigótico. En el protozoo mutante, la inversión de la asimetría viene determinada por un alelo situado en el locus denominado *janus* (en honor del dios romano de doble rostro, protector de las puertas de las ciudades). En las leguminosas son varios los alelos localizados en el locus que determina la dirección de enrollamiento de la legumbre; el alelo para el "enrollamiento-dextrógiro" es el dominante.

No nos sorprende que el principal modelo relativo a la herencia de las asimetrías lo ideara una estudiosa de la asimetría del desarrollo más intensamente investigada y menos entendida: la tendencia diestra o zurda en el hombre. Nos referimos a Marian Annett, psicóloga británica. Propuso que, mientras la preferencia de la mano derecha está asociada con un alelo "determinante de la derecha", no hay un alelo "determinante de la izquierda", sino tan sólo un alelo neutro que no determina dominancia alguna. Por tanto, cuando ambos progenitores tienen los dos alelos neutros, la tendencia no está

genéticamente determinada, y sus descendientes tienen la misma probabilidad de ser zurdos que diestros.

M. J. Morgan, del University College de Londres, y Michael C. Corballis, de la Universidad de Auckland, psicólogos a su vez, han ido más lejos: en su opinión todos los modelos de herencia de asimetrías deben ser como el de Annett, es decir, que un alelo codifica la direccionalidad y el otro alelo es neutro. Llegan a sugerir que los propios alelos codificados no determinan intrínsecamente la direccionalidad, sino que actúan sobre un gradiente de asimetría preexistente en el huevo fecundado.

El modelo alternativo al de Annett sugiere, en efecto, que los alelos codificadores "conocen la diferencia" entre izquierda y derecha. Vale decir: cuando el alelo del correspondiente locus codifica "derecha", resulta una asimetría derecha y cuando codifica "izquierda", la asimetría es zurda. En cualquiera de los modelos podría hallarse implicado más de un locus; pero ningún modelo especifica un mecanismo detallado del modo de acción.



APAREAMIENTOS EXPERIMENTALES de dos modelos hipotéticos que pretenden explicar la diferencia en la proporción de peces de ojos levógiros y de ojos dextrógiros de las poblaciones de platija estrellada de California y Japón. Al aparear una platija japonesa hembra de ojos levógiros (n.º 60) con dos de los machos japoneses de ojos levógiros (n.os 59 y 62), la descendencia de ambos apareamientos fue, en un 85 por ciento por término medio, de ojos levógiros. Aunque no representa la proporción típica de la población natural, es suficientemente elevada para eliminar la hipótesis no genética. Al aparear dos peces californianos de ojos levógiros (n.os 97 y 98) con dos de ojos dextrógiros (91 y 95), la descendencia dio, respectivamente, un 83 y un 18 por ciento de ojos levógiros. Este resultado también excluye la hipótesis no genética y contradice además la predicción de una de las hipótesis genéticas: que la descendencia de todos los acoplamientos entre platijas californianas sería la mitad de ojos levógiros y la mitad dextrógira.

La platija es asimétrica sólo en su estado adulto. La larva de platija recién eclosionada mantiene la simetría. En *The Life Story of the Fish: His Manners and Morals* el biólogo Brian Curtis cita a un periodista anónimo a propósito de esta metamorfosis. La narración, un tanto fantástica, se desarrolla como sigue: "La platija es el colmo ictiológico de la pereza. Comienza su vida en posición derecha, como cualquier otro pez; sin embargo, antes de alcanzar varias semanas de edad, comienza a cansarse de la lucha cósmica por la existencia. Se va hacia el fondo, se tumba de lado y rehúsa levantarse de nuevo. En esa posición se encuentra a sí mismo con un ojo mirando inútilmente hacia el fango. El ojo, aparentemente cansado de su esfuerzo por perforar el limo primordial, se comporta de modo aún no explicado por la ciencia. Se desplaza de un lado a otro y se une al otro ojo, que la fortuna ha colocado en la parte superior. De todo ello resulta uno de los peces de aspecto más estúpido, que, no obstante, ha alcanzado su objetivo. En ictínea indolencia se recuesta sobre el fondo, mal situado el ojo, atisbando hacia arriba por cualquier alimento que se le ponga a tiro. Lo que termina también por cansarle; cuando ello ocurre se entierra en el fango, donde ni siquiera tiene que mirar."

Para estudiar la genética de la asimetría de las platijas estrelladas tuve que establecer un programa de cruzamientos y criar la descendencia hasta la etapa en que sus ojos comienzan a migrar. Los huevos de la especie tienen algo menos de un milímetro de diámetro; a medida que el embrión se va desarrollando, ha de enroscarse apretadamente para caber en una esfera tan pequeña. La larva eclosiona rompiendo la membrana del huevo y coleteando libremente. En esa etapa apenas si es más que un saco de vitelo, dos ojos y una cola. El saco se reabsorbe en pocos días. La larva ha crecido ya de unos 2,2 milímetros hasta los 3,5 milímetros. Su boca se ha desarrollado y es capaz de alimentarse.

Nos hallamos ante una fase crítica de su crecimiento, porque, al agotarse el suministro de vitelo, la larva morirá de hambre si no empieza pronto a alimentarse. Hay que proporcionarle alimento que sea a la vez nutritivo y desmenuzado lo suficiente para que el alevín pueda ingerirlo. Sólo desde hace unos 20 años disponen los criadores de peces del alimento adecuado; se trata de un pequeño organismo marino, un rotífero del género *Brachionus*, que se cultiva fácilmente en grandes cantidades.

En cuanto los peces empiezan a procurarse el sustento, se desvanece la principal dificultad de su cultivo. Se alargan y engordan a la vez que se desarrollan sus aletas, huesos y pigmentación. El intervalo que media hasta el comienzo de la metamorfosis varía ampliamente, oscilando de 27 a 104 días tras la fecundación del huevo. Como factor de control importa más el tamaño que el tiempo transcurrido; con independencia de la edad, los peces miden unos siete milímetros de largo al inicio de la metamorfosis.

El proceso comienza cuando los huesos no se han osificado del todo y requiere la rotación del neurocráneo, que afecta al cerebro y a las órbitas de los ojos, pero no implica directamente al resto de la cabeza. Hay cierta reabsorción del hueso blando, que de otro modo impediría el desplazamiento del ojo;

aun así, el proceso es, fundamentalmente, migratorio, más que perforante. En la mayoría de las especies examinadas, incluida la platija estrellada, la aleta dorsal permanece detrás de la cabeza mientras los ojos migran. Después, la aleta crece hacia delante, hasta casi la punta del hocico. En contadas especies crece hacia delante antes de que el ojo haya terminado su desplazamiento; el ojo debe entonces moverse a través de la piel.

Para algunos especialistas, la migración del ojo hacia el lado incorrecto obedece a la tendencia que muestra el pez joven a inclinarse hacia un lado u otro o a apoyarse en el fondo del mar de un lado u otro. No parece que sea este el caso. Entre las miles de platijas que he observado a través de la metamorfosis no parecía haber relación entre el primer lado que escogen para posarse y el lado en el que terminaron los ojos. Es más, en algunos casos, el pez que está sufriendo la metamorfosis continuaba nadando en una posición derecha cerca del borde superior del acuario después de haberse desplazado el ojo migrante todo el trayecto hasta la parte superior de su cabeza.

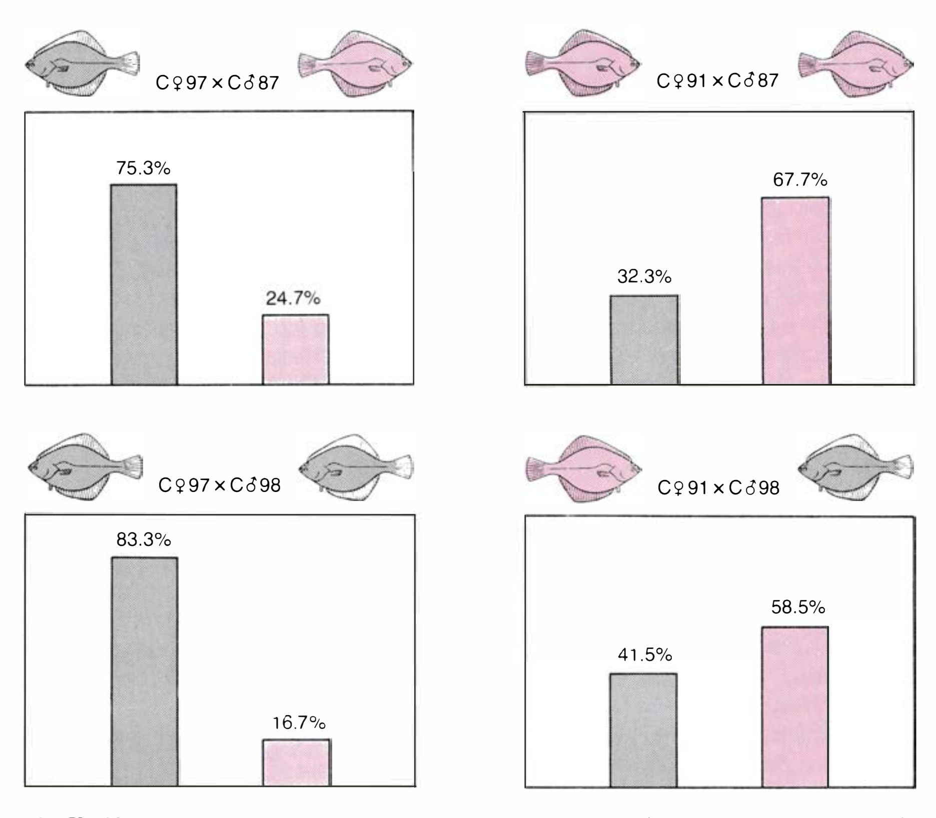
Las platijas estrelladas son los peces planos más adecuados donde estudiar la herencia de las asimetrías. Por tres razones. Primero, la diferencia en el porcentaje de peces de ojos levógiros es tan alta (desde 50 por ciento en California a casi 100 por ciento en Japón) que es fácil decidir estadísticamente si la proporción de peces de ojos levógiros en la descendencia de un cruzamiento se aparta de la proporción de la población parental. Segundo, la existencia misma de esta clina implica verosímilmente la acción de algún mecanismo de selección natural. Tercero, en ausencia de datos experimentales, los peces parecen amoldarse al modelo genético que denominaré "M & C". Lo he llamado así, entre comillas, por la razón siguiente: no obstante su semejanza con la versión del modelo de Annett propuesto por Morgan y Corballis, se trata de un modelo general cuya aplicación no se circunscribió específicamente a las platijas.

De acuerdo con el modelo "M & C", se espera que las platijas estrelladas japonesas sean homocigóticas para los alelos de ojos levógiros. Todas ellas reciben un alelo de ojo levógiro de cada progenitor. En la platija estrellada de California se espera que sean homocigóticas para alelos neutros. Predice este modelo que cualquier cruzamiento que se dé entre dos peces japoneses producirá, sobre todo, descendencia de ojos levógiros, mientras que el cruzamiento entre dos peces californianos producirá igual número de descendientes de ojos levógiros que de ojos dextrógiros.

Una hipótesis alternativa propone que en la población japonesa sólo hay alelos de ojos levógiros, mientras que en la población californiana aparecen alelos de ojos levógiros y de ojos dextrógiros, de forma que algunos peces californianos serían homocigóticos para alelos de ojos dextrógiros, otros serían homocigóticos para alelos de ojos levógiros y unos terceros serían heterocigóticos. Sin embargo, la población japonesa sería homocigótica para alelos de ojos levógiros. De esta situación puede predecirse, como antes, que cualquier cruzamiento entre peces japoneses producirá principalmente descendencia de ojos levógiros. En cambio, de dos cruzamientos entre dos pares distintos de peces californianos pueden esperarse diferentes porcentajes de descendencia de ojos levógiros, ya que los alelos que porten los peces californianos seguramente serán diferentes. (Esta hipótesis incluye subhipótesis sobre tipos de herencia y frecuencias aléticas de uno u otro ojo.)

Según otra hipótesis no hay diferencia genética entre platijas estrelladas de ojos levógiros o dextrógiros; y ello vale sólo para los peces de ojos levógiros y dextrógiros de California (esta parte de la hipótesis concuerda con "M & C") y vale también para los peces de ojos levógiros y ojos dextrógiros tanto californianos como japoneses (esta parte no concuerda con "M & C"). La diferencia en la proporción de peces de ojos levógiros entre las poblaciones de Japón y California se debería a algún factor ambiental. De esta hipótesis no genética se desprende que, si las descendencias crecieran en ambientes idénticos, todos los posibles cruzamientos producirían el mismo porcentaje de peces de ojos levógiros.

A fin de someter a prueba las predicciones realicé varios cruzamientos entre un total de 13 platijas estrelladas: cuatro hembras (una de Japón y tres de California) y nueve machos (tres de Japón y seis de California). Una hembra y tres machos (californianos todos) eran de ojos dextrógiros. Cultivé la descendencia en pequeños estanques del New England Acuarium, de Boston. El agua salada que llenaba los es-

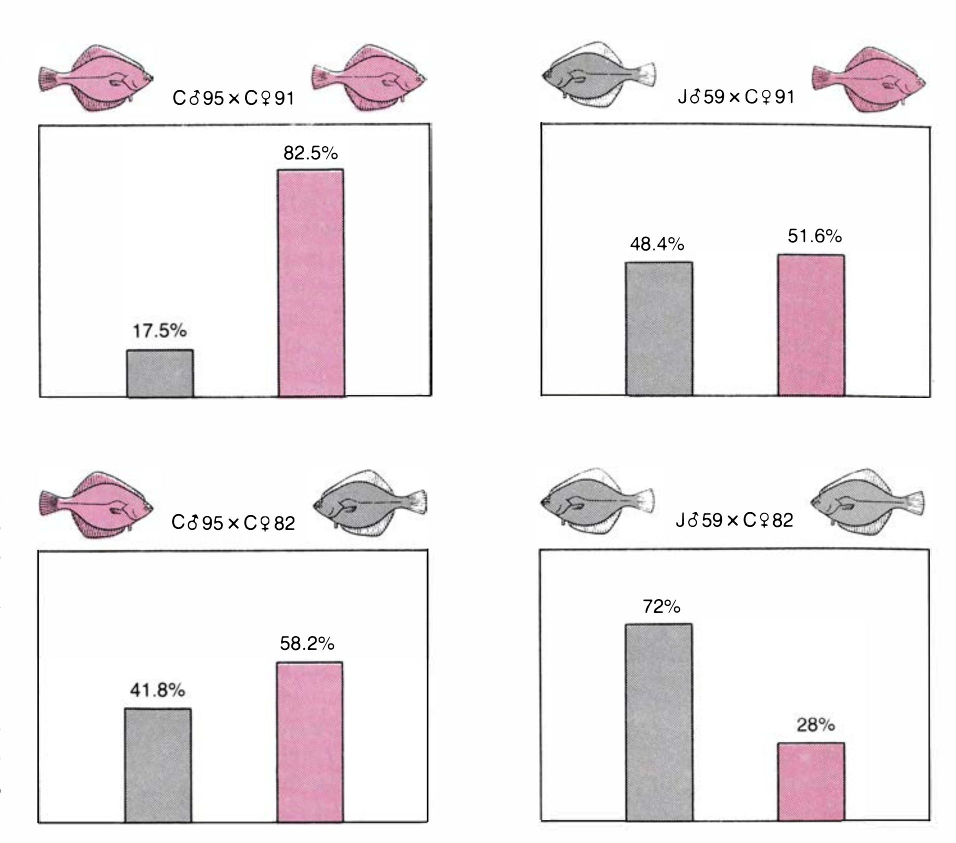


INFLUENCIA DE LA HEMBRA en la proporción de descendencia de ojos levógiros. Dos hembras californianas la prueban. Al cruzar una hembra de ojos levógiros (n.° 97) con un macho californiano de ojos dextrógiros (n.° 87), la descendencia era en un 75 por ciento de ojos levógiros. Cuando una hembra de ojos dextrógiros (n.° 91) se apareó con ese macho, sólo el 32 por ciento de la progenie era levógira. Al cruzar las hembras con un macho californiano de ojos levógiros (n.° 98), la hembra de ojos levógiros dio un 83 por ciento de la descendencia de ojos levógiros y, la hembra de ojos dextrógiros, menos del 42.

tanques estaba en todos a la misma temperatura y procedía de una fuente común. Todas las larvas recibían el mismo alimento. Con estas medidas se aseguraba un ambiente idéntico para todas las descendencias.

Los resultados del cultivo apoyaban la primera de las dos hipótesis alternativas al modelo "M & C" y rechazaban las predicciones de la hipótesis del modelo "M & C" y de la hipótesis no genética. A saber:

Los cruzamientos entre la única hembra japonesa y dos de los machos japoneses produjeron dos grupos de descendientes con una proporción media del 85 por ciento de ojos levógiros. Dos cruzamientos californianos distintos (uno entre una hembra de ojos dextrógiros y un macho de ojos dextrógiros y el otro entre una hembra de ojos levógiros y un macho de ojos levógiros) produjeron descendencias que fueron, respectivamente, de sólo el 18 por ciento de ojos levógiros y del 83 por ciento de ojos levógiros. Estos resultados de California van contra la predicción "M & C", que afirma que toda la descendencia de los peces de California será en un 50 por ciento de ojos levógiros. Por supuesto, tanto los resultados de los cruzamientos japoneses como los de los californianos contradicen la hipótesis no genética.



INFLUENCIA DEL MACHO en las asimetrías de la descendencia; la demostraron dos machos de platija estrellada, uno de California y otro del Japón. Cruzando el macho de platija californiana de ojos dextrógiros (n.º 95) con una hembra californiana de ojos dextrógiros (n.º 91), menos del 18 por ciento de la progenie era de ojos levógiros. El apareamiento entre el macho de ojos levógiros del Japón (n.º 59) y la misma hembra produjo una descendencia con más del 48 por ciento de ojos levógiros. Al aparear los machos con una hembra de ojos levógiros californiana (n.º 82), el californiano dio una descendencia con un 42 por ciento de ojos levógiros y, el japonés, una descendencia en un 72 por ciento levógira.

Los resultados de otros cruzamientos señalaban en algunos peces cierta tendencia hacia la producción de descendencia, bien de ojos levógiros, bien de ojos dextrógiros. Una hembra californiana de ojos levógiros, aun apareándola con un macho californiano de ojos dextrógiros, produjo más del 75 por ciento de ojos levógiros en su descendencia. (Esa misma hembra produjo casi un 85 por ciento de su descendencia de ojos levógiros, cuando, como se ha dicho, se la apareó con un macho californiano de ojos levógiros.) Sin embargo, al aparear la única hembra californiana de ojos dextrógiros con los dos mismos machos de California mencionados más arriba, se produjo, con el macho de ojos dextrógiros, un 32 por ciento de descendencia de ojos dextrógiros y, un 41 por ciento, con el macho de ojos levógiros. De un modo similar, uno de los machos japoneses de ojos levógiros, apareado con una hembra californiana de ojos levógiros, dio una descendencia con un 72 por ciento de ojos levógiros y, con la hembra californiana de ojos dextrógiros, una descendencia con un 46 por ciento de ojos levógiros.

Probablemente la proporción de descendencia de ojos levógitos y dextrógiros dependa, hasta cierto punto del ambiente. Los dos cruzamientos entre padres japoneses produjeron, respectivamente, descendencias con el 13 y el 17 por ciento de ojos dextrógiros. En Japón, las platijas estrelladas de ojos dextrógiros constituyen menos del uno por ciento de la población. Quizá la dotación genética de la única hembra japonesa fuese anormal, lo que explicaría por qué se produjeron tantos descendientes de ojos dextrógiros. No obstante, yo me inclino por una explicación ambiental. Muchas de las larvas de las platijas presentaban anomalías en su desarrollo. Casos hubo que no mostraron migración ocular. Es probable, pues, que el ambiente artificial de los estanques no igualara las condiciones del ambiente natural necesarias para un proceso de desarrollo normal, incluida la determinación de la asimetría.

Estos resultados acentúan también una diferencia entre las asimetrías de los caracoles y las platijas. En los cruzamientos entre platijas estrelladas, ambos progenitores, macho y hembra, influyen en la proporción de la descendencia de ojos levógiros y dextrógiros. En los caracoles, sólo el genotipo de la hembra (su combinación de alelos) influye en el sentido de enrollamiento de las conchas de la progenie. Además, toda la progenie de una hembra se enrolla en la misma dirección. Los huevos fecundados de la mayoría de los moluscos se desarrollan según un proceso denominado de segmentación espiral; los huevos de los vertebrados se desarrollan por segmentación radial. "Espiral" y "radial" aluden a las primeras divisiones celulares de los huevos fecundados. En los moluscos, los planos de esas primeras divisiones están girados entre sí, de lo que resulta un desarrollo en espiral del embrión joven. En los vertebrados, los patrones de segmentación varían, mas nunca en dirección espiral.

Según los experimentos realizados por C. P. Raven, de la Universidad de Utrecht, que trabajó con el caracol de enrollamiento dextrógiro *Lymnaea*, la dirección de segmentación y la de enrollamiento están determinadas por una asimetría en el citoplasma del huevo. Raven observó que la pauta de segmentación de los embriones que darían caracoles de enrollamiento levógiro era la imagen especular de la pauta de seg-

mentación de los de enrollamiento dextrógiro. Determinó que tales asimetrías derivaban de asimetrías preexistentes en el huevo antes de su fecundación. Esto es, la asimetría de la estructura del huevo se establecía en él cuando todavía formaba parte del cuerpo de la hembra. De ahí que sea el genotipo de la hembra y no el del macho el que determine la dirección de enrollamiento de la descendencia. Los huevos de las platijas, por supuesto, no tienen segmentación espiral, como tampoco es exclusivamente materno el control del sentido de sus asimetrías.

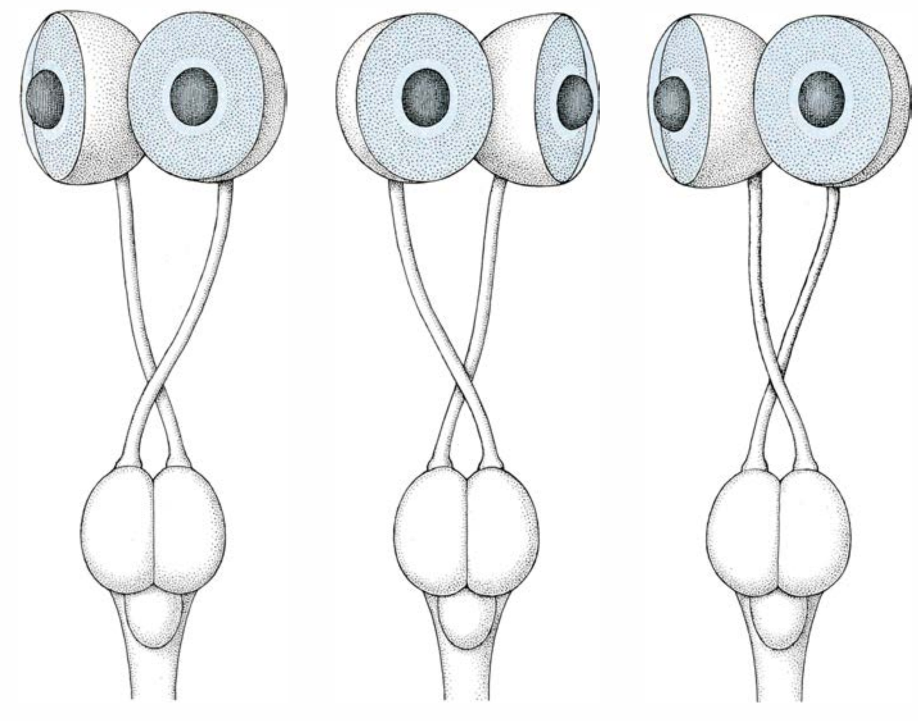
No puede descartarse que la diferencia entre las proporciones de platijas estrelladas de ojos levógiros y dextrógiros de California y del Japón se deba a distintas respuestas del desarrollo frente a ambientes diferentes. La división entre peces planos de ojos dextrógiros y levógiros es claramente hereditaria. Ello plantea la cuestión de si la tendencia a tener los ojos en uno u otro lado presenta algún significado adaptativo. Intuitivamente parece ventajoso que una platija concreta tenga la misma lateralidad, izquierda o derecha, que los demás miembros de su población. Se apoya ese razonamiento intuitivo en que son muy escasas las especies de peces planos que muestran una lateralidad repartida a partes iguales. También lo avala la conducta de apareamiento de los peces. Si el macho monta la hembra, como ocurre en algunas especies, son de imaginar las dificultades que se le presentarán a un pez de ojos dextrógiros que trate de aparearse con un ejemplar de ojos levógiros. Cualquier traba al apareamiento llevaría a la especie a desarrollar una sola clase de lateralidad.

Más fácil resulta pensar que la tendencia a presentar ojos levógiros constituye una dificultad para la platija plateada. Los peces planos, la platija estrellada incluida, tienen asimetría interna, que afecta también a los nervios ópticos. Como en todos los peces óseos, los nervios ópticos se cruzan: el nervio óptico derecho se une al lado izquierdo del cerebro, y viceversa. En los mamíferos se da cierto intercambio de fibras nerviosas a nivel del quiasma (el cruzamiento de las fibras), lo que a duras penas ocurre en los peces; de ahí que pueda distinguirse sin ambigüedad un nervio óptico derecho y otro izquierdo.

A comienzos de siglo, G. H. Parker, de la Universidad de Harvard, publicó un estudio de los quiasmas de peces óseos. Descubrió que en los peces simétricos, así el bacalao, el nervio óptico izquierdo se cruzaba por encima del nervio óptico derecho en la mitad de la población, y que el nervio óptico derecho hacía lo propio sobre el izquierdo en la otra mitad. Algunos peces planos remedaban la condición del bacalao en este punto; pero otros no: el nervio del ojo que habitualmente emigra era el que estaba por encima. La platija plateada pertenece al segundo grupo. Parker consideró que dicha disposición debíase a que con ello se facilitaba la emigración del ojo. Razonó que, si el ojo izquierdo migraba cuando el nervio óptico derecho estaba encima, los nervios se retorcerían, mientras que si el nervio óptico izquierdo estaba en la posición más alta, no habría retorcimiento. Sugirió que el retorcimiento constituiría una desventaja desde el punto de vista mecánico, lo que explicaría la evolución seguida por la disposición del nervio óptico en los peces planos.

Si ello fuera cierto, seguiríase que todas las platijas estrelladas japonesas, de ojos levógiros y con su nervio óptico izquierdo más alto que el derecho, que es el que migra, se encontrarían en desventaja. Cuesta aceptar que tal situación sea fruto de la selección natural, como difícil resulta imaginar cualquier tipo de asimetría ambiental que confiriese ventaja adaptativa a la preferencia por ojos levógiros. Por un lado, no es fácil idear asimetrías ambientales que el pez no pudiera invertir con sólo darse la vuelta. Por otro, aun cuando admitiésemos que una hipotética asimetría de ese tipo causara la predominancia de ojos levógiros en la población de platijas estrelladas de aguas japonesas, ¿cómo explicar entonces la presencia de muchas especies de peces planos que portan ojos dextrógiros en las mismas aguas?

La única conclusión razonable parece ser la que niega que exista diferencia adaptativa entre ojos levógiros y ojos dextrógiros y afirma que los dos caracteres deben estar asociados genéticamente con algún otro factor, todavía no identificado, que sí la presenta. En el caso de que tal conclusión fuera cierta, la situación valdría de ejemplo de una dificultad con que topan frecuentemente los biólogos que estudian la evolución: la de decidir qué características son adaptativas y cuáles son selectivamente neutras. Por el momento, en lo que respecta a la diferencia entre especies de peces planos de ojos levógiros y dextrógiros, esa diferencia, por sorprendente que resulte, parece ser un señuelo evolutivo.



LOS NERVIOS OPTICOS de las platijas, como los de todos los peces óseos, se cruzan de tal forma que el nervio del ojo izquierdo va al lóbulo derecho del cerebro y el nervio del ojo derecho se dirige al lóbulo izquierdo. Se muestran en la ilustración tres casos de tales cruzamientos. A la izquierda se distingue el quiasma (el cruzamiento de las fibras nerviosas) de un fletán de California, tras haber migrado su ojo derecho al lado izquierdo de la cabeza. En medio se observa el quiasma de una platija estrellada de ojos dextrógiros una vez que su ojo izquierdo se ha situado ya en el lado derecho de la cabeza. En la derecha se observa el quiasma de una platija estrellada de ojos levógiros después de haber migrado su ojo derecho: se aprecia un "retorcimiento" de los nervios ópticos. Este retorcimiento puede constituir una desventaja y sugiere que, cualquiera que sea la razón de la predominancia de ojos levógiros en la población japonesa de platijas estrelladas, no puede considerarse un tipo de adaptación selectiva para ojos levógiros.

Chorros cósmicos

La violenta actividad que registra el centro de muchas galaxias se manifiesta en la producción de estrechos haces fuertemente colimados de plasma (gas ionizado). Los hay cuya longitud multiplica un millón de veces el año-luz

Roger D. Blandford, Mitchell C. Begelman y Martin J. Rees

Las observaciones astronómicas hechas con radiotelescopios han revelado que el centro de muchas galaxias es un lugar de violenta actividad, que suele manifestarse, como acaba de descubrirse, mediante la producción de chorros cósmicos. Por tales hemos de entender estrechos haces de plasma (gas ionizado) que parecen salir lanzados desde el centro de una galaxia, emitiendo radioondas al mismo tiempo. Cada chorro puede adquirir una longitud superior a un millón de años-luz. Termina en una extensa mancha de radioemisión situada bastante lejos de la imagen óptica de la galaxia. La energía contenida en esta mancha puede exceder 10^{60} erg, cantidad que viene a ser equivalente a la conversión en energía de toda la masa aportada por 10 millones de estrellas. La observación óptica hace tiempo había dado cuenta ya de la existencia de algunos de estos chorros; pero hubo que esperar hasta años muy recientes para que las nuevas técnicas de radioastronomía detectaran la frecuencia de los chorros en el universo. Vale decir: los chorros deben producirse bajo condiciones muy variadas; mas no hay acuerdo sobre el modo preciso en que se producen.

Los chorros cósmicos pueden adquirir muchas formas distintas. Uno de los ejemplos mejor estudiados es el chorro asociado a la galaxia elíptica NGC 6251, que se halla a 300 millones de años-luz de nuestra propia galaxia. En 1977, Peter Waggett, Peter Warner y John Baldwin, de la Universidad de Cambridge, descubrieron un chorro, largo y estrecho, que emanaba del núcleo de NGC 6251. El chorro tiene una anchura angular de sólo tres grados, lo que no obsta para que su longitud supere los 400.000 años-luz. Su extremo angosto coincide con el mismo centro de la galaxia; los observadores de Cambridge hallaron allí una pequeña fuente de radioemisión, casi puntual. En 1978, Anthony C. S. Readhead y Marshall H. Cohen, del Instituto de Tecnología de California, pudieron cartografiar esta fuente puntual. Descubrieron que también ella estaba formada por un chorro estrecho con una fuente puntual en un extremo. Se diría que habían abierto una muñeca rusa cuyo interior sólo ofreciera una versión más reducida de la anterior. El chorro pequeño, colineal con el mayor, no mide nada más que tres años-luz de longitud.

De aspecto algo distinto es el chorro 3C 449, que está asociado a una galaxia elíptica distante 100 millones de años-luz. En 1979, Richard A. Perley y sus colaboradores, trabajando con el radiotelescopio Gran Distribución, en Socorro, Nuevo México, hallaron que dos chorros emergían del centro de la galaxia en sentidos opuestos. Ambos chorros mostraban varios ángulos o codos bien marcados. El chorro septentrional se doblaba bruscamente hacia el Este a unos 100.000 años-luz del centro; volvíase a doblar luego y seguía su rumbo hacia el Norte hasta que parecía doblarse de nuevo. El chorro meridional también se doblaba hacia el Este y lo hacía luego hacia el Sur. Hay varios chorros que presentan este comportamiento en zig-zag.

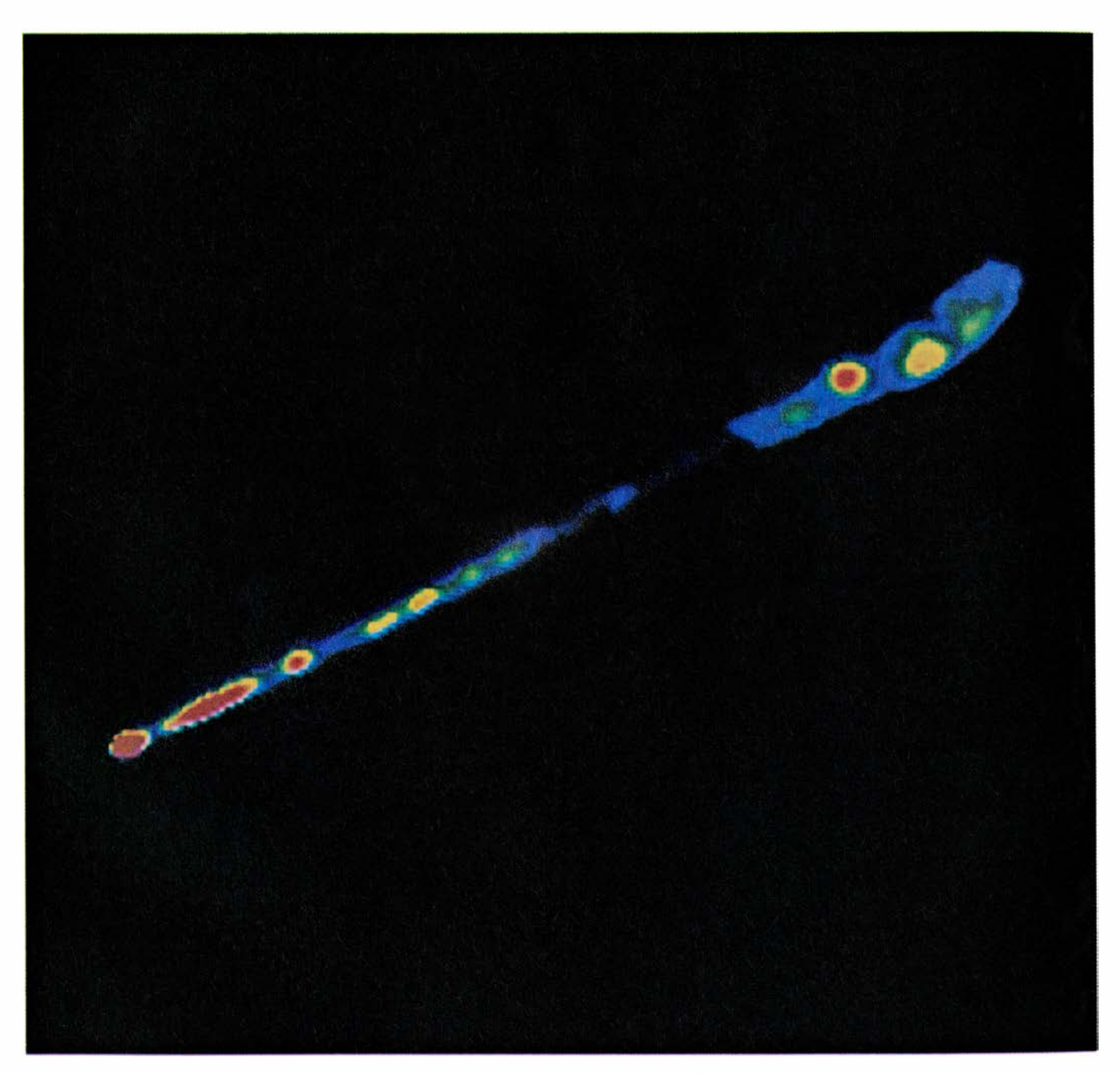
El fenómeno de los chorros no está confinado a otras galaxias. Tenemos un buen ejemplar en la nuestra propia, a sólo 15.000 años-luz del Sol. Se llama SS433. Del análisis de las observaciones ópticas se deduce que dos chorros emergen en sentidos opuestos de un sistema de estrella binaria a una velocidad de 80.000 kilómetros por segundo, algo más de la cuarta parte de la velocidad de la luz. Se han detectado ya esos chorros en las longitudes de onda de radio y en las de rayos X. La naturaleza se las arregla para producir chorros dentro de objetos ligeros, cuya masa multiplica un número de veces muy bajo la del Sol, y en objetos pesados, de hasta mil millones de soles (masa del núcleo de una galaxia). Algo parecido cabe afirmar de la longitud de los chorros: los hay que no cubren siquiera la distancia entre la Tierra y el Sol, pero también los hay que la superan 10.000 millones de veces.

Avanzamos antes que los chorros no constituían ninguna novedad para los astrónomos. En 1917, Heber D. Curtis, del Observatorio Lick, descubrió un chorro óptico asociado a la gran galaxia elíptica M87, que se encuentra en un cúmulo rico de galaxias en la constelación Virgo. Pero hubo que esperar hasta 1953 para que R. C. Jennison y M. K. Das Gupta, de los Labortorios Nuffield de Radioastronomía, construyeran el primer radiointerferómetro: dos radiotelescopios conectados eléctricamente de suerte que registren detalles de la radiofuente cuya mínima entidad impide su detección individual por cada radiotelescopio. Jennison y Das Gupta apuntaron su interferómetro hacia la radiofuente del Cisne A, cuya asociación con una galaxia elíptica distante habían puesto de manifiesto antes Walter Baade y Rudolph Minkowski, de los Observatorios de los Montes Wilson y Palomar. Descubrieron, con gran sorpresa, que la radioemisión no emanaba de la propia galaxia, sino que procedía de una mancha difusa a cada lado de la misma. Los radiotelescopios que se vienen empleando ahora presentan creciente complicación y potencia, lo que ha permitido mostrar que la mayoría de las fuentes extragalácticas detectadas a radiofrecuencias de menos de 1 gigahertz (1000 millones de ciclos por segundo) tienen esta doble estructura básica. Desde 1953, dar con la razón de ello se ha constituido en auténtico reto para los profesionales de la astronomía extragaláctica. Como veremos, el descubrimiento de los radio-chorros condujo a las puertas de la solución.

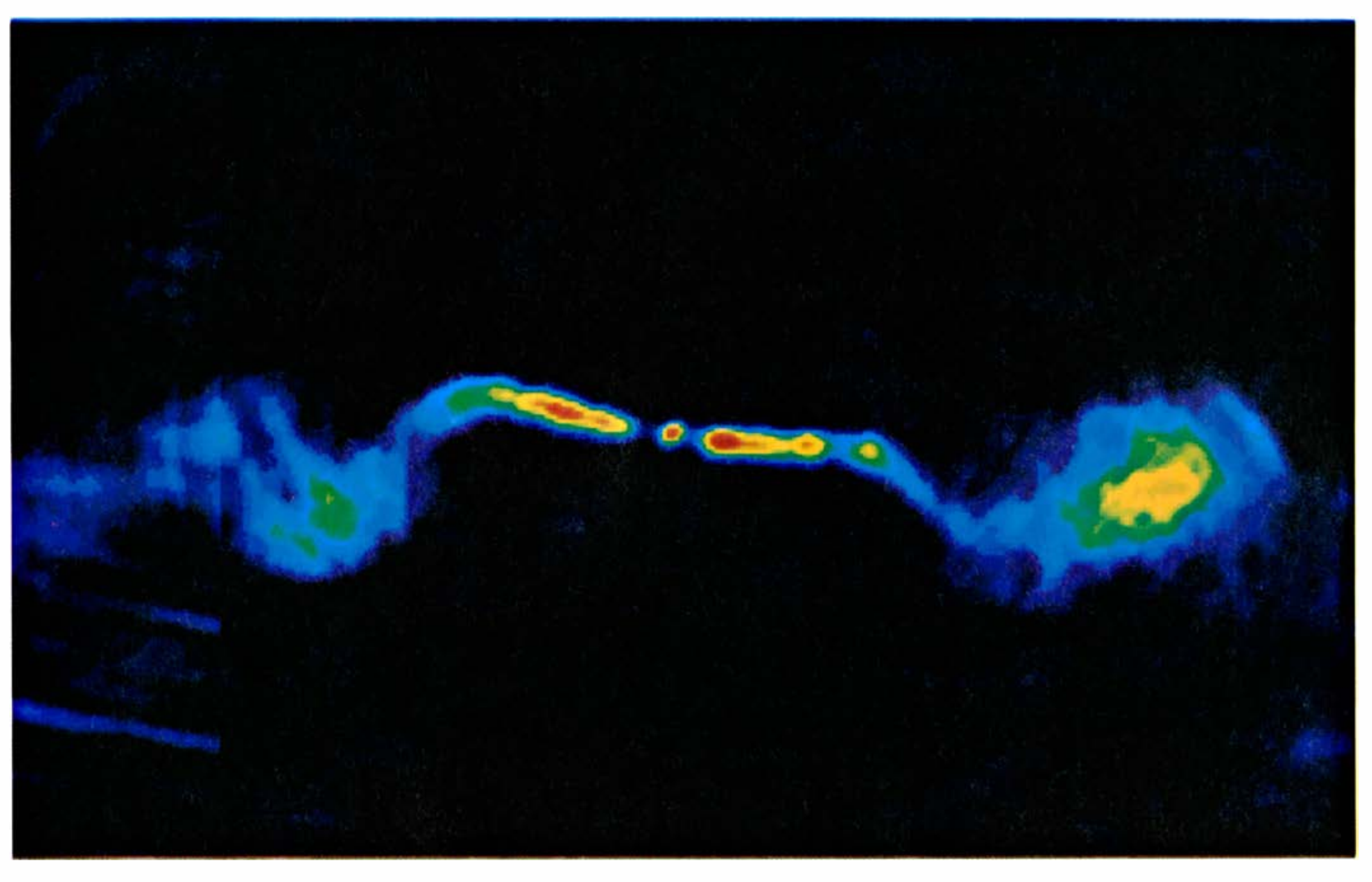
Hacia el final del decenio de 1960 y durante el de 1970, se construyó una se-

rie de interferómetros cada vez mayores, de los que mencionaremos el de la Universidad de Cambridge, Observatorio Westerbork en Holanda y Observatorio Nacional de Radioastronomía (NRAO) en Virginia Occidental. El empeño más reciente es la Gran Distribución de Nuevo México. Consta de 27 radiotelescopios conectados, cada uno de los cuales tiene 25 metros de diámetro. Puede revelar características que subtiendan un ángulo en el cielo no mayor de dos décimas de segundo de arco. (Este es el ángulo subtendido por una moneda de 10 centavos a una distancia de 20 kilómetros.) En 1971, George Miley y Campbell M. Wade usaron el interferómetro NRAO para desvelar la presencia de manchas calientes en Cisne A. Cada mancha es una región que emite intensa radiación a longitudes de onda de radio. Se encuentra en la extremidad externa de cada lóbulo, o mancha difusa, de las radiofuentes dobles más brillantes (como la del Cisne A). Son bastantes las radiofuentes dobles más brillantes que muestran colas o puentes de baja intensidad de emisión que, arrancando de las manchas calientes, se extienden hacia atrás, camino del centro de la fuente, donde suele haber un foco compacto de emisión llamado núcleo. La mayoría de las radiofuentes dobles (piénsese en Cisne A) están asociadas a una galaxia elíptica o un quasar (un objeto brillante y puntual fuera de la galaxia); en tales casos, el núcleo se encuentra invariablemente en el centro óptico de la fuente.

La resolución de los mapas construidos con radiointerferómetros queda limitada por la separación de los distintos radiotelescopios: cuanto mayor la separación, tanto mayor la resolución. En la Gran Distribución, los telescopios distan hasta 32 kilómetros



CHORRO DE UNA SOLA RAMA y de unos 400.000 años-luz de longitud, que se encuentra a 300 millones de años-luz de nuestra galaxia. Es el chorro más rectilíneo que se conoce. Su origen (*parte inferior izquierda*) coincide con el centro de la galaxia elíptica NGC 6251. Se sabe que su origen coincide con un chorro de sólo tres años-luz que está alineado con el chorro más largo. La radioemisión de NGC 6251 y de las radiogalaxias que aparecen en las páginas 76 y 77 se cartografió en falso color con la ayuda de los 27 radiotelescopios que constituyen la Gran Distribución (VLA) cerca de Socorro (Nuevo México).



CHORROS CON SIMETRIA ESPECULAR en la radiofuente 3C 449. Emergen en direcciones opuestas de una galaxia elíptica. Cada chorro se dobla hacia arriba y después hacia abajo. Vuelven luego a aparecer en un gran lóbulo de radioemisión. Se cree que los chorros transportan energía del centro de la galaxia a los lóbulos. Los codos con simetría especular pueden deberse a que la galaxia esté describiendo una órbita en torno a una compañera.

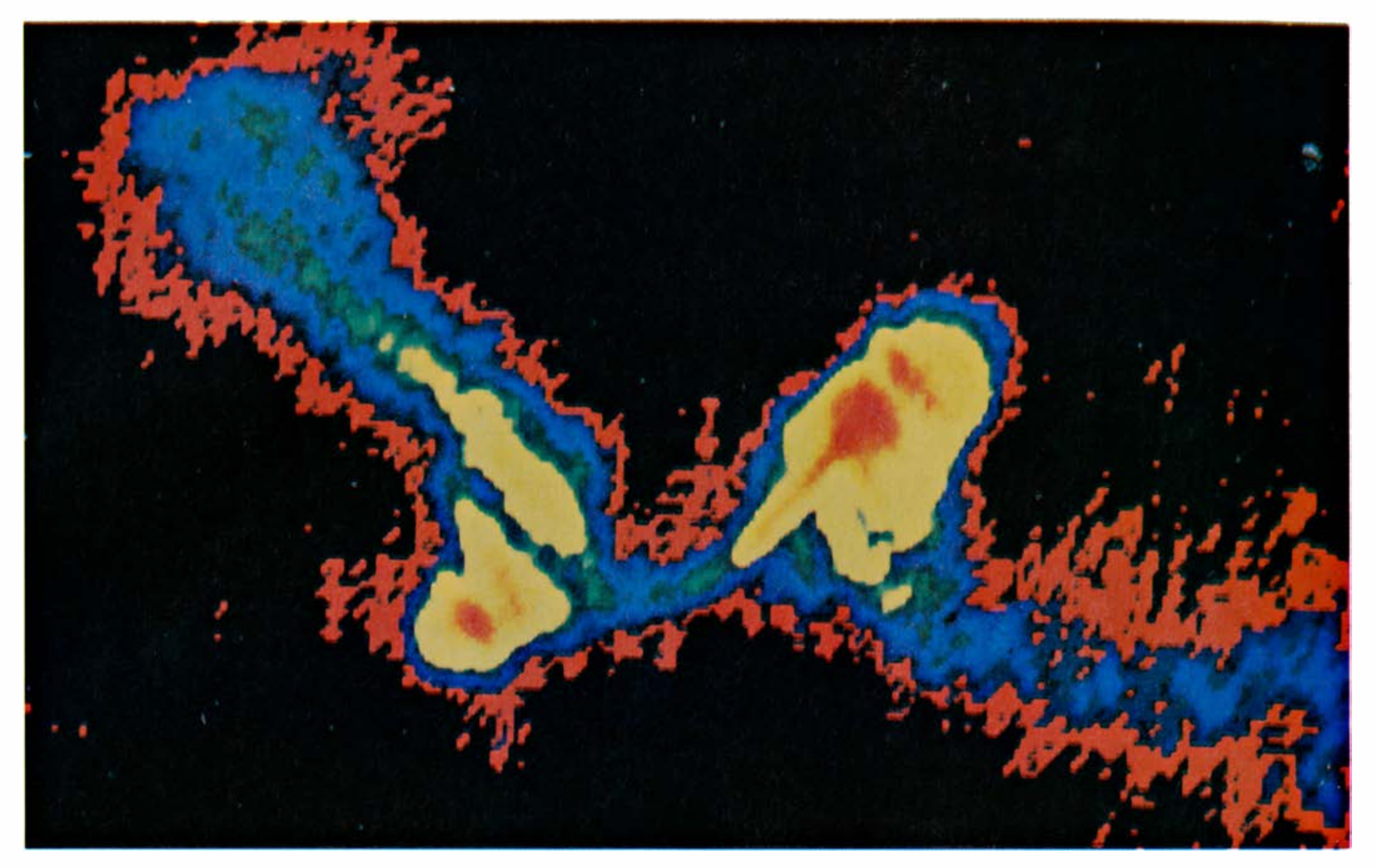
unos de otros. En la técnica llamada interferometría de base muy larga (o VLBI), que se desarrolló simultáneamente con la construcción de la moderna generación de radiointerferómetros, los telescopios pueden hallarse en continentes distintos. La resolución que se alcanza así es proporcionalmente mayor. Se pueden distinguir detalles de hasta una milésima de segundo de arco, es decir, el ángulo subtendido por una moneda de diez centavos a 4000 kilómetros, por un año-luz a la distancia de la galaxia NGC 6251 y por 100 años-luz a la distancia de los quasars más lejanos.

Tamaña resolución exige su tributo. Los distintos telescopios no pueden enlazarse directamente, lo que obliga a registrar en cinta magnética las radioseñales de los telescopios y compararlas bastante después de realizadas las observaciones. A causa de la pérdida de información que ello implica, los mapas construidos con interferometría de base muy larga carecen de la sensibilidad y el detalle (en la escala de las observaciones) que presentan los hechos con interferometría de base más corta. Sin embargo, los astrónomos han cosechado éxitos recientes en el uso regular de cuatro o más radiotelescopios en trabajo de muy larga base; combinaciones que han permitido avanzar grandes pasos en el camino de la mejora de los resultados. Los mapas VLBI que se están construyendo hoy no tienen nada que envidiar a los que se hacían con telescopios enlazados 10 años atrás.

La interferometría de muy larga base se ha apuntado sus buenos triunfos en la exploración de los núcleos de las radiofuentes. Ha mostrado que generalmente presentan detalles que no se han podido resolver. Estos rasgos han de poseer, por consiguiente, un tamaño angular inferior a algunas milésimas de segundo de arco. En contraste, las manchas calientes de las fuentes suelen carecer de detalles en pequeña escala. La VLBI ha tenido éxito también en el estudio de radiofuentes compactas, por lo regular asociadas ópticamente con quasars distantes. En cada una de ellas, el flujo de ondas de radio procede generalmente de un núcleo pequeño, no de un lóbulo amplio. No obstante, se detecta normalmente radioemisión a un nivel más bajo, que procede de una región (típicamente de forma irregular) en torno al núcleo. Se ha hallado que las radiofuentes compactas suelen variar de intensidad en escalas de tiempo que oscilan desde meses hasta años.

Se han encontrado chorros con interferómetros enlazados y con los de base muy larga. Y también mediante una tercera técnica, puesta a punto en Jodrell Bank, en la que los radiotelescopios están enlazados por transmisión en microondas. Se sabe que en más de 70 fuentes dobles extragalácticas emanan chorros en gran escala que salen del centro de la galaxia asociada; los núcleos de seis de esas dobles fuentes, al menos, despiden chorros en pequeña escala. Aun cuando todavía no se ha explorado como sería menester una muestra suficiente de radiofuentes, han aparecido ya algunas indicaciones útiles.

Por un lado, la probabilidad de hallar chorros en gran escala en radiofuentes dobles de baja potencia es mayor que la de encontrarlos en radiofuentes dobles de alta potencia. Obedece esto, en parte, a que la mayoría de las fuentes potentes están a grandes distancias, lo que dificulta la detección de sus chorros; pero no basta con acogerse a la distancia para dar razón completa de esa tendencia. Piénsese que no se han observado todavía chorros en gran escala en Cisne A, una potente fuente doble que está relativamente próxima en



CHORROS CON SIMETRIA INVERSA en la radiofuente 4C 26.03. También emergen en direcciones opuestas de una galaxia elíptica (NGC 326). En este caso, un chorro se dobla hacia arriba y el otro lo hace hacia abajo. De nuevo, sin embargo, cada chorro termina en un extenso radiolóbulo. Los codos con simetría de inversión pueden resultar de la precesión de la fuente de donde brotan los chorros, esto es, el giro del eje de rotación de la fuente.

lo que se refiere a tales objetos. (Dista de nosotros unos 450 millones de años-luz.) Cuando de una radiofuente potente emana un chorro en gran escala, éste suele dirigirse hacia un solo lado, sin que se encuentre ningún contra-chorro. Los chorros asociados a radiofuentes más débiles son generalmente dobles (se extienden hacia ambos lados).

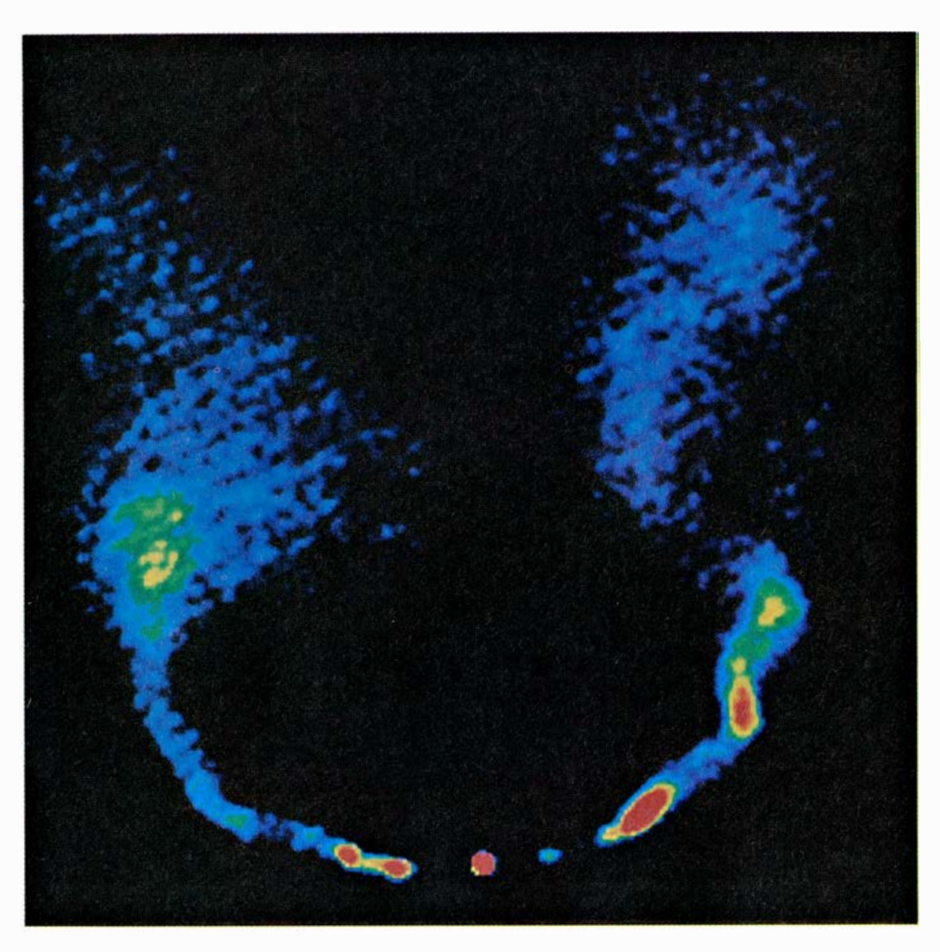
Conocemos una segunda tendencia, que se funda en la observación siguiente: la radioemisión procedente de fuentes extragalácticas está invariablemente polarizada, esto es, los campos eléctricos en las radioondas están preferentemente dirigidos a lo largo de una línea en el cielo. Se sabe la razón de ello. Se da por supuesto que la radioemisión se genera mediante el proceso de sincrotrón, en el que los electrones que se mueven con velocidades próximas a la de la luz son acelerados por un campo magnético y emiten así ondas electromagnéticas. Los electrones son acelerados en una dirección perpendicular a la orientación del campo; por consiguiente, la radiación de sincrotón estará polarizada en esa dirección.

Se deduce de ello que, cartografiando la configuración de la polarización para una radiofuente dada, podemos inferir la orientación del campo magnético de la fuente. Además, si medimos la intensidad de la polarización, sabremos si la geometría del campo magnético está muy ordenada, o si lo está poco. Resulta que en chorros potentes, las líneas de fuerza del campo magnético guardan una ordenación paralela a la dirección del chorro. En chorros débiles, las líneas tienden a ser perpendiculares. En chorros de intensidad intermedia, se observa normalmente una región de transición desde un campo paralelo en el chorro cerca del centro de la galaxia asociada hasta un campo perpendicular mucho más lejos chorro adelante.

Antes incluso de que se hubieran hallado buenos ejemplos de radiochorros en radiofuentes dobles, cierto número de teóricos (los autores incluidos) habían argumentado que los lóbulos de una radiofuente doble deberían hallarse recibiendo un suministro continuo de energía. Y como la emisión procedente del núcleo de la radiofuente doble era clara prueba de la incesante actividad del centro de la galaxia a la que estaba asociada la fuente doble, parecía deducirse la existencia de un enlace energético entre el centro y los extensos radiolóbulos. La indicación apuntaba sin ambigüedades hacia una idea: la de que los chorros eran corrientes de gas que salían del centro de una galaxia. Los chorros proveerían entonces a los lóbulos no sólo de energía, sino también de masa, cantidad de movimiento y flujo magnético.

Iniciada la actividad en el centro, el chorro se propaga hacia fuera atravesando la galaxia, de forma semejante al chorro de agua que sale con fuerza de la boca de una manguera. El chorro cósmico cruza primero el medio interestelar y luego el intergaláctico. La densidad de materia varía allí aproximadamente desde un átomo de hidrógeno por centímetro cúbico a uno por cada millón de centímetros cúbicos. A pesar de ello, el chorro debe apartar en su avance la materia que le sale al paso; de aquí que la punta del chorro corra más lenta que el gas que fluye dentro del haz. ¿Qué ocurre? La energía se acumula en la punta del chorro; tal es la interpretación plausible de las manchas calientes.

Se cree que el flujo de gas es supersónico, es decir, la velocidad del gas es mayor que la de una onda sonora en el gas. Sin embargo, cuando éste se acerca a la mancha caliente, frena bruscamente. Ello provoca la formación de



RADIO-COLAS de NGC 1265, una galaxia elíptica en un rico cúmulo de galaxias. Tal vez representen chorros arrastrados en una larga curva por el gas intergaláctico que sopla junto a la galaxia a una velocidad de varios miles de kilómetros por segundo. El mapa se realizó, en falso color, con el VLA.

una onda de choque a través de la sección del chorro. El efecto del choque resultante es considerable. Antes de que el gas alcance esa onda, la mayor parte de su energía se encuentra en la forma de energía cinética ordenada. El paso a través del frente de onda convierte una porción sustancial de esta energía ordenada en dos formas: energía de los electrones relativistas (electrones que se mueven a velocidades cercanas a la de la luz) y energía de un campo magnético. Resulta, pues, natural del todo que la mayor parte de la radioemisión intensa en una radiofuente doble se origine por la acción del gas circundante donde decelere el chorro.

Tras haberse detenido en las manchas calientes, la materia del chorro fluye otra vez hacia la galaxia. Así expande los radiolóbulos que se ven en los radiomapas. La corriente de retorno arrastra consigo los electrones relativistas y las líneas del campo magnético. Cualquier partícula del chorro que consideremos pasará un tiempo bastante corto (aproximadamente de 10.000 a un millón de años) en una mancha caliente. La mayor parte de la energía de la radiofuente se almacena, por tanto, en los lóbulos, que probablemente acumulan el gas durante 100 millones de años. El contenido de energía de los lóbulos es prodigioso. En una fuente como el Cisne A, debe ser al menos la cantidad que se liberaría si toda la masa de 100.000 estrellas se convirtiese en energía. Y en manantiales mayores, excederá la masa equivalente a un millón de estrellas.

Muchos chorros presentan grandes codos. En algunos casos, esa forma puede ser una manifestación de una inestabilidad que se produce en él a medida que va excavando su camino a través del medio intergaláctico. Cualquiera que haya manejado una manguera flexible de jardinería tiene experiencia del fenómeno. Si la manguera se deja en el suelo formando una curva y se abre el agua, la manguera comienza a temblar. La inestabilidad del flujo de agua de la manguera obedece a que el agua de la parte exterior del codo circula ligeramente más deprisa para ir al paso del agua que fluye por la parte interior del codo. Consiguientemente, la presión en la parte exterior se hace menor que la presión en la parte interior; la configuración de la manguera cambia, aumentando su curvatura. En sus líneas generales, el fenómeno físico es el mismo para los chorros cósmicos, aunque los detalles se complican mucho más cuando se trata de una corriente supersónica.

Mas no todos los codos se producen al azar en el chorro cósmico. Consideremos las fuentes llamadas radiocolas. A baja resolución, la radioemisión de las mismas no parece extenderse a horcajadas de una galaxia óptica, como ocurre cuando se trata de una fuente doble típica, sino que se extiende en una larga curva a un lado de la galaxia. A alta resolución, la diferencia queda explicada. Se ven dos chorros que emergen del centro de la galaxia; mas no terminan en manchas calientes, sino que se doblan continuamente, de suerte que ambos se unen con la larga y curvada cola de la radioemisión.

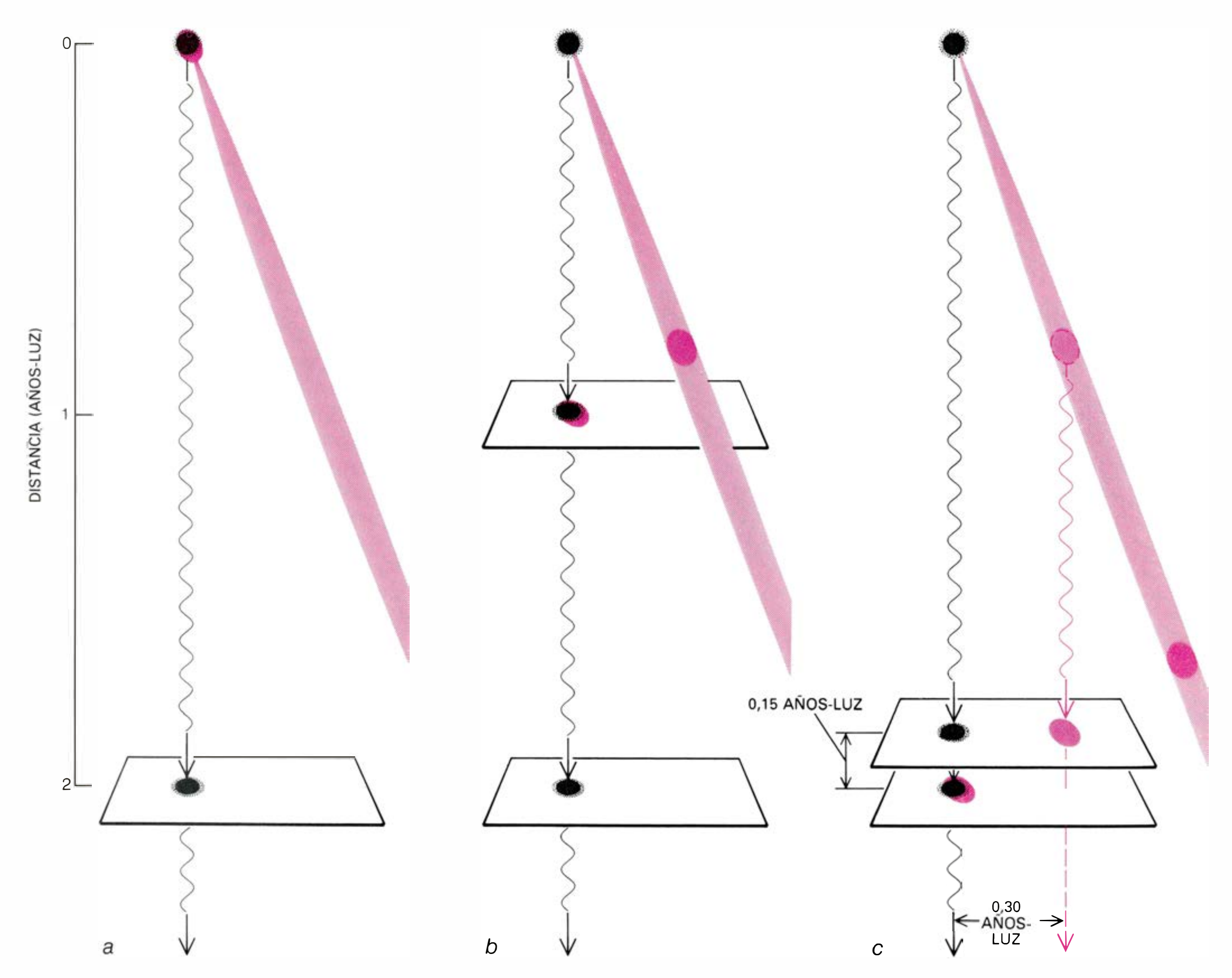
Por lo que se aprecia, lo que ocurre en las radiocolas es lo siguiente: en el centro de la galaxia se forman dos chorros de potencia bastante baja que son arrastrados lateralmente por el medio intergaláctico, al fluir junto a la galaxia con velocidades de varios miles de kilómetros por segundo. El fenómeno recuerda al humo que sale de una chimenea y es arrastrado por un fuerte viento. El hecho de que las radiocolas se encuentran preferentemente en cúmulos ricos de galaxias apoya esta interpretación de su morfología. (El espacio que media entre galaxias de los cúmulos ricos se cree está lleno de gas ionizado y caliente bastante denso.) Merece destacarse de las radiocolas el hecho de que, al parecer, sus chorros pueden doblarse en ángulo casi recto por el viento intergaláctico sin perder su identidad. Queda así claro que los chorros no siempre son inestables. Se presume que una vez unidos los chorros a la cola, su gas es impulsado hasta una velocidad comparable con la del medio intergaláctico. Si atendemos entonces a la velocidad del medio con respecto a la galaxia y recordamos la longitud de la cola, deduciremos la edad de la radiofuente. En un caso típico, la edad resulta ser de 300 millones de años.

Se ha propuesto una explicación que guarda parentesco con la anterior para fuentes del tipo a que pertenece 3C 449, cuyos chorros muestran varios codos notables, pero son lineales en sus grandes rasgos. En general, estas fuentes no se encuentran en los cúmulos ri-

cos de galaxias. Por otra parte, suelen tener galaxias compañeras cercanas. Es probable, por tanto, que la radiofuente corresponda a una galaxia en órbita alrededor de sus compañeras. Si la velocidad del gas en los chorros no es mucho mayor que la velocidad orbital, el movimiento de la galaxia quedará "escrito" en el cielo en forma de codos de los chorros.

En cierto sentido, los codos de los chorros constituyen una ilusión: el gas no fluye describiendo tales codos. Imagínese una fotografía de alguien que está regando un jardín balanceando una manguera. El agua que queda "pendiente" en el aire presenta una configuración en zig-zag, pero cada gota de esa configuración se mueve como si se tratara de un proyectil balístico. La forma de los codos de un chorro cósmico sirve para deducir las posiciones tridimensionales de las galaxias a partir de sus proyecciones bidimensionales en el cielo. Aun así, los chorros podrían concordar con la presencia de una galaxia en órbita en el caso exclusivo de que los codos de un chorro se correspondan con los codos del otro, a modo de imagen especular. Si un chorro se dobla hacia la derecha en el cielo, el otro deberá doblarse hacia la derecha a la misma distancia de la galaxia, más o menos. La simetría especular requerida parece que se da, realmente, en fuentes tales como 3C 449.

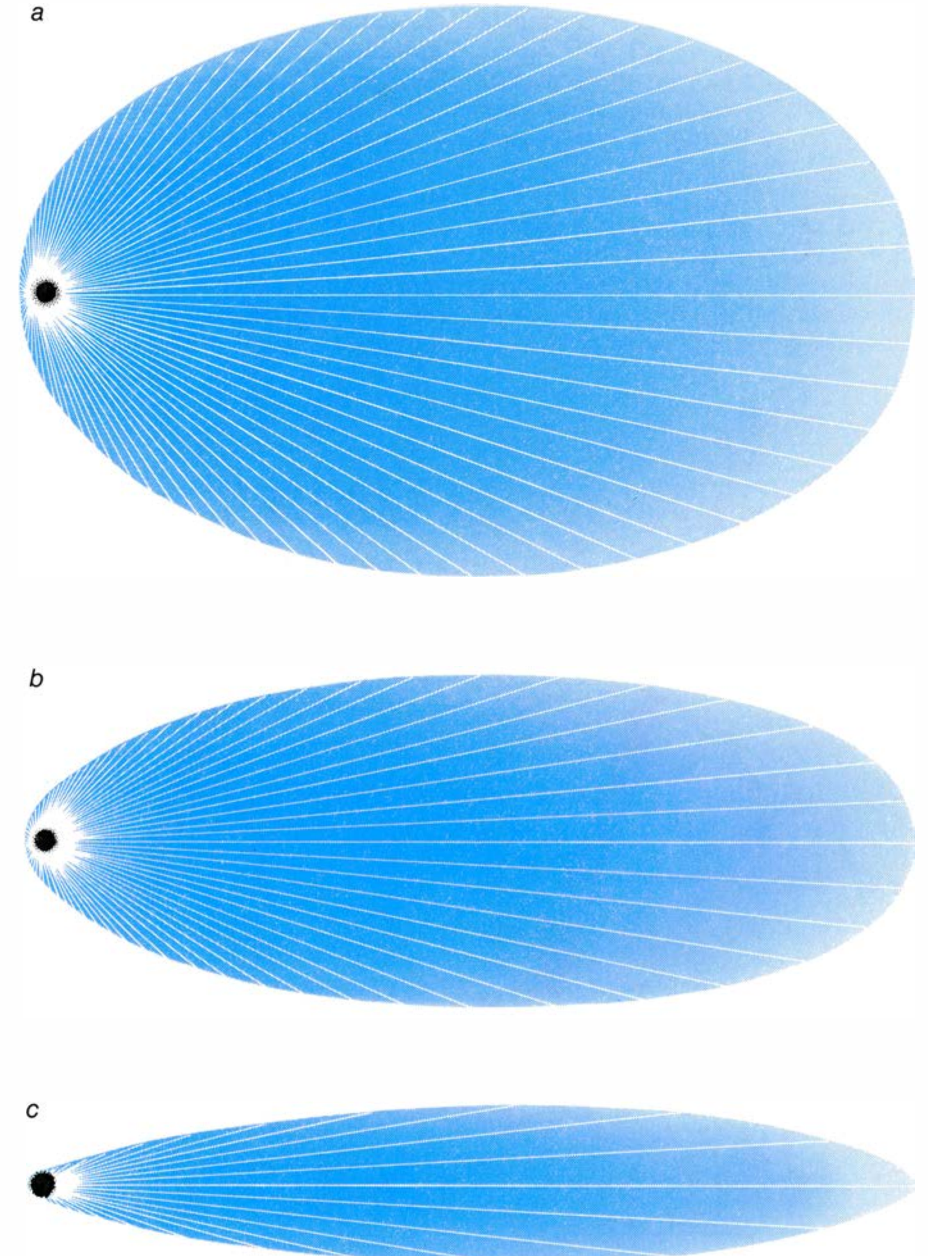
Una clase de radiofuentes distinta muestra simetría inversa. En este caso, un codo hacia la derecha de un chorro corresponde a un codo hacia la izquierda del otro. El mejor ejemplo de chorros con simetría inversa es SS433. Se cree que la fuente de los chorros tiene aquí movimiento de precesión, como un trompo. De nuevo, las partículas de gas que hay en esos haces avanzan en línea recta, pero la figura que se observaría en un instante particular semejaría la trayectoria del sacacorchos. Recientes radioobservaciones realizadas con la Gran Distribución por Robert M. Hjellming, del Observatorio Nacional de Radioastronomía, y Kenneth J. Johnston, del Laboratorio de Investigación de la Armada de los Estados Unidos, parecen evidenciar eso, precisamente. Quizá los chorros extragalácticos que presentan simetría inversa proceden de una fuente dotada de movimiento de precesión, como SS433, pero que está situada en el centro de una galaxia. Tal precesión constituiría



SE CREE QUE LA EXPANSION SUPERLUMINICA (movimiento aparente de componentes de una radiofuente a velocidades que exceden la de la luz) es mera ilusión. Sin embargo sugiere que tales componentes tienen velocidades próximas a la de la luz. En este ejemplo, un observador situado a dos años-luz de una radiofuente (*a*) no se da cuenta de que la fuente ha emitido en su chorro un brillante nudo de gas ionizado (*en color*). Un año después (*b*), el observador ve todavía una fuente única; la luz que muestra la aparición del nudo ha de recorrer un año-luz para llegar a él. Mientras tanto, el nudo se ha movido con respecto a la fuente en una dirección que forma un ángulo de 20 grados con la visual. Su velocidad es 0,9 veces la lumínica. Transcurridos dos años (*c*), el observador distingue el nudo. La luz del nudo emitido en *b* sólo está 0,15 años-luz detrás. De aquí que, en 0,15 años-luz, el observador apreciará que el nudo se ha movido hasta 0,3 años-luz de la fuente. Y podría deducir que el nudo se está moviendo con una velocidad doble de la de la luz.

una pista importante para descubrir qué condiciones del centro conducen a la producción de chorros.

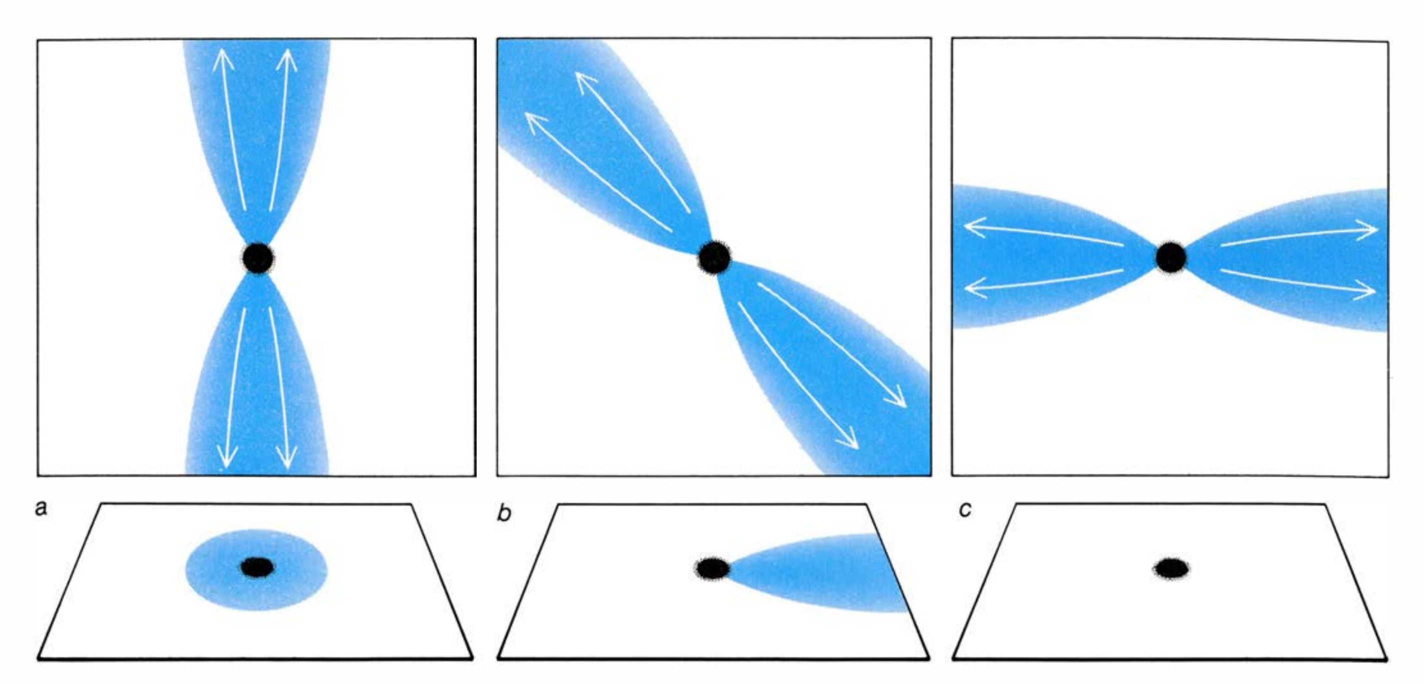
Vayamos ahora a los pequeños chorros descubiertos por interferometría de muy larga base. A lo largo de los últimos tres años, se han cartografiado varias radiofuentes compactas y varios núcleos compactos asociados a radiofuentes extensas con detalle suficiente para revelar un chorro en pequeña escala. En la mayoría de los casos, el chorro presenta una morfología característica: un chorro único que emerge de un punto no resuelto de intensa radioemisión. En muchos ejemplos, el chorro en cuestión apunta en la dirección aproximada de uno de los radiolóbulos de la fuente. Si hay un chorro no detectado que apunta en la dirección opuesta, debe ser al menos diez veces más débil que el observado. Readhead, Cohen y sus colegas, del Instituto de Tecnología de California, han descubierto una importante correlación que tiene que ver con la dirección de los chorros pequeños. Los asociados a núcleos compactos de radiofuentes extensas (es decir, las fuentes cuya emisión proviene principalmente de dobles lóbulos) tienden a guardar una buena alineación con los lóbulos. La falta de alineación no supera algunos grados. En contraste, los chorros de pequeña escala asociados a radiofuentes compactas ofrecen una curvatura apreciable y así están desalineados con la débil estructura de la fuente, típicamente en ángulos de 30 grados o más.



LA ABERRACION RELATIVISTA focaliza la radiación emitida por un objeto que se mueve a una velocidad próxima a la de la luz, de manera que el objeto radia intensamente en la dirección de su movimiento. En *a*, el emisor (una nube de gas ionizado que intrínsecamente radia por igual en todas las direcciones) se mueve hacia la derecha a la mitad de la velocidad de la luz. En *b* avanza a 0,75 veces la velocidad de la luz, en *c* a 0,94 veces la velocidad de la luz y en *d* a 0,98 veces la velocidad de la luz. El emisor se ha hecho invisible salvo para quien se halle delante de él. La forma de cada configuración muestra la manera como la intensidad de la radiación varía con el ángulo de emisión. Visto directamente, de frente, el emisor en *a* es varias veces más brillante que una fuente estacionaria que radiase por igual en todas las direcciones, en *b* es 30 veces más brillante, el emisor *c* lo es 440 veces y en *d* 3100 veces.

A la hora de interpretar esta correlación y dar una explicación de alcance general para las radiofuentes, hemos de señalar dos fenómenos cruciales. El primero, llamado expansión superlumínica, ha quedado bastante convincentemente demostrado en recientes mapas VLBI, aunque se sospechaba de su existencia desde hace diez años. La expansión superlumínica significa, a grandes trazos, que hay rasgos de ciertos chorros VLBI que se mueven en el cielo con respecto a la posición del núcleo no resuelto de donde parte el chorro, según se ha observado. Se diría que el rasgo se mueve a una velocidad superior a la de la luz. El caso más convincente de expansión superlumínica ha sido documentado por Timothy J. Pearson y sus colegas del Cal Tech. Se refiere al quasar 3C 273, cuya fama entre los astrónomos le viene de su chorro óptico de 60.000 años-luz de longitud. El núcleo de 3C 273 emite un chorro VLBI, que según la generalización de Cohen y Readhead, está doblado y desalineado con respecto al chorro óptico. A lo largo de los últimos cuatro años, una brillante mancha en el chorro VLBI ha aumentado su separación respecto al núcleo en casi un 50 por ciento. En la hipótesis de que 3C 273 se encuentre a 1500 millones de años-luz de nuestra galaxia, la mancha brillante se habría movido con una velocidad quíntuple de la de la luz.

Los astrónomos no afirman exactamente que la mancha avance más rápidamente que la luz. Semejante postula-



DIFERENTES ORIENTACIONES de los chorros cósmicos. Esas orientaciones distintas confieren a los haces una morfología diversa a causa de la aberración relativista. Aquí, cada radiofuente está formada por dos chorros que emergen en direcciones opuestas desde el centro de una galaxia. El centro radia por igual en todas direcciones, pero el gas de los chorros se mueve a velocidad suficiente para que los chorros radien preferentemente en la dirección del movimiento. Si los chorros son paralelos a la visual (*a*), un observador distante percibirá un núcleo de radioemisión rodeado de un radiohalo. Si los chorros son casi paralelos a la visual (*b*), detectará sólo el núcleo y el chorro más próximo; el más lejano será invisible, porque apenas si radiará hacia el observador. Si los chorros son perpendiculares a la visual (*c*), se distinguirá el núcleo porque la aberración tornará invisibles ambos chorros.

do contradiría la teoría especial de la relatividad; y nadie habla aquí de revisar, en sus cimientos, la física. Lo que denuncian es que somos víctimas de una ilusión. La mancha brillante parece correr a mayor velocidad que la luz porque la luz procedente de la mancha y la luz que viene del núcleo del quasar, que nos sirve de referencia, viajan a lo largo de caminos de diferente longitud y, por tanto, tardan en llegar a nosotros intervalos de tiempo ligeramente distintos. Cuál sea la configuración precisa sigue en el alero; mas la alternativa más sencilla requiere que la mancha brillante se esté moviendo bajo un pequeño ángulo con la visual entre la radiofuente y los observadores terrestres y con velocidad ligeramente inferior a la de la luz. Parecerá entonces que se mueve superlumínicamente. El nudo central del problema no es que la expansión superlumínica sugiera velocidades superiores a la de la luz, sino velocidades muy próximas a la de la luz para los componentes de las radiofuentes.

El segundo fenómeno a tener en cuenta antes de analizar las radiofuentes es la aberración relativista. Supóngase que un cazador desea disparar a un pato cuando éste se encuentra en su vertical. Habrá de apuntar su escopeta hacia arriba y disparar antes de que el pato pase por encima. Los perdigones se dirigirán verticalmente hacia arriba hasta el ave, que para entonces habrá avanzado y llegado al mismo tiempo que la perdigonada (o así lo espera el cazador). Consideremos ahora la situación desde la óptica del pato. Ve al cazador apuntándole; los perdigones no trazan una vertical perfecta, sino que tienen una componente horizontal en su movimiento. Dicho de otra manera, los perdigones viajan con una ligera desviación que inclina su trayectoria a lo largo de la dirección en que el cazador se está moviendo con respecto a la presa.

Lo mismo ocurre con los fotones o partículas de radiación electromagnética. De aquí que una nube de plasma que radie fotones por igual en todas las direcciones dará la impresión de estar brillando preferentemente a lo largo de su dirección de movimiento. Cuando el plasma se mueve a velocidad relativista (esto es, casi tan deprisa como los fotones), el efecto se acentúa grandemente. Tómese el caso del plasma cuyo movimiento relativo da la impresión de expansión superlumínica en una radiofuente. Aquí, aproximadamente la mitad de los fotones que el plasma emite se radiarán según un cono cuyo eje se prolongue en la dirección de movimiento del plasma y posea un ángulo de abertura de sólo cinco a 10 grados. Los fotones se habrán hecho más energéticos en virtud de un corrimiento Doppler que aumenta su longitud de onda. El resultado neto es sorprendente. Si un observador está en el cono, creerá que la fuente brilla entre 100 a 1000 veces más que una fuente estacionaria. Pero si el observador se halla fuera del cono, y alejado lo suficiente, no distinguirá tal fuente.

Podemos adelantar ya una poderosa interpretación unificadora de las radiofuentes extragalácticas. Supóngase que la mayoría de las fuentes constan de un par de chorros relativistas que parten, en direcciones opuestas, desde el corazón del núcleo galáctico. Si los chorros están más o menos alineados con la dirección de la visual entre la fuente y la Tierra, detectaremos el chorro que apunta hacia nosotros, con mayor o menor enfoque. El otro chorro no se ve. La fuente aparece brillante y como si emitiera sólo hacia un lado. Los chorros acaban en un par de extensos radiolóbulos que radian fotones de forma casi igual en todas las direcciones. Producen el débil halo de emisión extensa que, como se observa, rodea a las fuentes compactas. No es de esperar que los chorros sean uniformes y rectos de un modo exacto. Presentarán, sin duda, codos y heterogeneidades naturales. En las radiofuentes compactas, los codos quedan exagerados porque nuestra visual guarda cierta alineación con los chorros. Si las heterogeneidades se mueven a lo largo del chorro a velocidades relativistas, nos dará la sensación de que se mueven superlumínicamente.

En esta interpretación, las radiofuen-

tes extensas se observan a grandes ángulos con respecto a la dirección de sus chorros. Para las más potentes de esas fuentes, como la Cisne A, los chorros son relativistas, e invisibles por tanto. Para fuentes algo menos potentes, piénsese en la radiofuente asociada a la galaxia NGC 6251, los chorros son débilmente relativistas. Sin embargo, basta la magnitud de la aberración de su radioemisión para que detectemos sólo el chorro que queda a nuestro lado del centro de la radiogalaxia. Para fuentes débiles, del tenor de la 3C 449, los chorros son subrelativistas; vale decir: ambos son patentes.

La interpretación de las radiofuentes extragalácticas que toma en consideración el efecto Doppler sobre los haces explica tal variedad de observaciones que resulta improbable que sea incorrecta en todos sus términos. Pero conviene señalar que la interpretación no es fácil de reconciliar con determinadas observaciones. Por ejemplo, la densidad del gas en algunos de los chorros a un solo lado puede adquirir un valor muy elevado, tanto que si su velocidad fuera relativista la potencia transportada por el chorro sería 10.000 veces mayor que la potencia aparentemente disipada por el chorro y los radiolóbulos de la fuente.

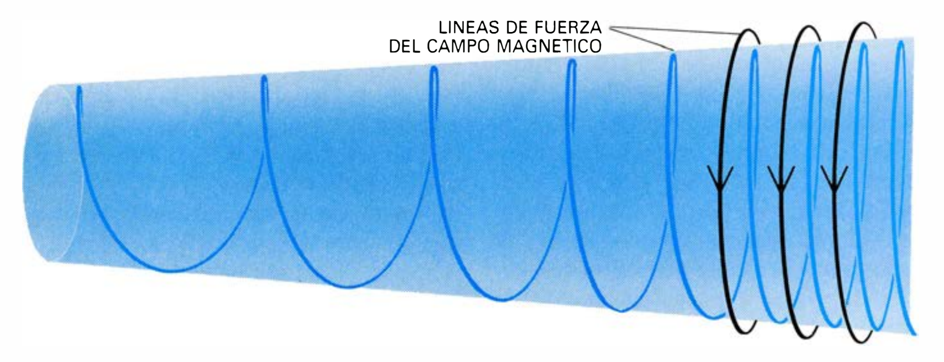
Parece, pues, que, en algunas radiofuentes, el chorro a un solo lado no se sigue del movimiento relativista. Una hipótesis que reemplaza a la descrita propone que la visibilidad de un chorro tenga que ver con su estabilidad. En las fuentes más intensas, los chorros serían conducciones estables e invisibles que transportarían eficazmente energía desde el núcleo hasta los radiolóbulos. Los chorros podrían, sin embargo, convertirse en inestables. La inestabilidad canalizaría presto la energía cinética de los chorros transformándose en aceleración de los electrones relativistas y en una radioemisividad fuertemente intensificada. Los chorros a un solo lado se producirían cuando solamente uno de ellos se tornara inestable. Los chorros a ambos lados se hallarían asociados a fuentes de baja potencia, donde el transporte de energía es poco eficaz. Una prueba de observación que podrá dilucidar cuál de las explicaciones contrapuestas de los chorros a un solo lado en radiofuentes es la correcta será determinar si el lóbulo asociado al chorro en esta clase de fuentes es sistemáticamente diferente del lóbulo que parece carecer del chorro. Si esto ocurre, costará defender la hipótesis de la influencia del efecto Doppler sobre los haces.

Una dificultad que mina la investigación de chorros cósmicos es que no se conoce la cantidad total de potencia que radian los chorros. Sabemos cuánto radian en longitudes de onda de radio; nada más. El problema es que los radiotelescopios son mucho más sensibles que los telescopios ópticos. Se han detectado algunos chorros, no muchos, en longitudes de onda ópticas o de rayos X. Esto significa que su potencia óptica debe centuplicar, por lo menos, su potencia en radio. No hay que excluir la posibilidad de que las radiofuentes presenten su máxima luminosidad a frecuencias ópticas.

No acaban ahí las dificultades. Cuesta dirimir si un chorro delineado por radioemisión es o no idéntico a la región donde el gas ionizado está fluyendo. Si una gran parte de la verdadera sección recta del chorro es invisible, magnitudes tales como la potencia y el impulso del chorro tenderían a subestimarse. Los cálculos realizados presentan un interesante rompecabezas. Algunos de los chorros más luminosos parecen colimarse a medida que se alejan de la galaxia emisora. ¿Qué es, entonces, lo que estrecha el chorro? ¿La presión del gas que lo rodea en el medio intergaláctico? En ese caso, el calentamiento y los choques producidos en el gas convertirían la región circundante del chorro en una potente fuente de rayos X. ¿Son anillos de líneas del campo magnético que el chorro transporta consigo? Propiedad de las líneas de fuerza del campo magnético cuando adquieren una configuración curva en el chorro es que ejercen cierta tensión; un anillo del campo magnético que arrolle así al chorro actuaría como una banda de goma. Desconcierta, por otro lado, que los chorros más intensos, aquellos que necesitan más su confinamiento, posean campos magnéticos que tienden a ser paralelos, y no perpendiculares, al eje del chorro. Pero la dificultad se disiparía si sólo el eje del chorro es lo bastante luminoso para ser observable.

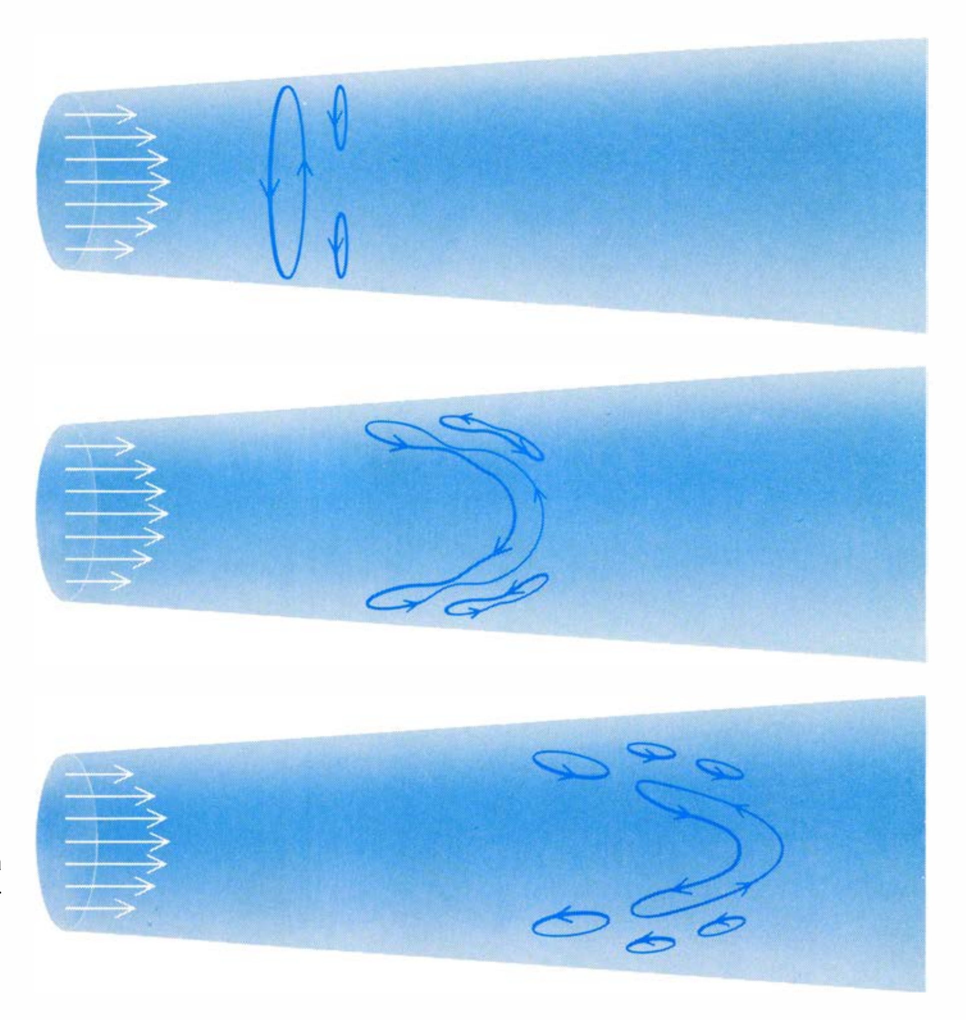
Para el astrofísico teórico, la pregunta central en el tema que aborda el artículo es la referente a la causa de los chorros. Han aparecido un sinfín de esquemas, pero ninguno goza de apoyo unánime. Pudiera producirlos una boquilla. Imagínese una copiosa provisión de gas caliente en la parte central de una gran nube, más fría. El gas caliente puede explicarse por la presencia de objetos astronómicos infrecuentes, tales como agujeros negros o púlsares que calientan el gas cuando éste cae en su intenso campo gravitatorio. Imagínese que la nube fría está en rotación. El gas caliente será menos denso y tenderá a escapar por la línea de menor resistencia, dirigida a lo largo del eje de rotación de la nube. La nube le ofrece sus paredes flexibles al gas que sale; a medida que huye, su presión disminuirá para hacerse igual a la presión externa. Lo que puede terminar en un chorro crecientemente bien colimado.

El mecanismo hipotético de la boquilla contradice las observaciones realizadas en varias radiofuentes. Por un lado, parece probable que los chorros VLBI surjan en regiones de tamaños no superiores a algunos años-luz. Pero si una nube de gas mostrara la suficiente densidad para confinar la presión de un chorro en lugar tan pequeño, constituiría una espectacular fuente de rayos X. Además, el alineamiento entre el chorro VLBI y los chorros en gran escala de las radiofuentes extensas sugiere que el mecanismo colimador debe actuar como un buen giróscopo: tendrá larga "memoria" para una dirección fija en el espacio. Y no está claro que una nube gaseosa sea un buen giróscopo.



LINEAS DE FUERZA DEL CAMPO MAGNETICO en un chorro. Estas responden a la dilatación del chorro poniéndose más perpendiculares a la dirección de la corriente en el chorro de gas. Básicamente, el campo del chorro tiene una componente paralela a la dirección de la corriente y otra componente perpendicular a ella. La dilatación del chorro hace crecer la primera más que la segunda. Así, los bucles de una línea de fuerza helicoidal (que representa la superposición de un campo paralelo y otro perpendicular) puede inclinarse hacia la perpendicular al aumentar la distancia a la radiofuente. Los anillos perpendiculares del campo pueden ejercer una tensión sobre el chorro y detener de esta forma su dilatación.

Esas dificultades pierden peso si se supone que la colimación se produce en la región que rodea un agujero negro cuya masa puede tener cualquier valor entre un millón de masas solares y varios miles de millones. La presencia de un tal agujero negro en el centro de las radiogalaxias viene apoyada por varias pruebas circunstanciales independientes. En primer lugar, algunas de las radiofuentes más activas varían su flujo óptico de un día a otro, lo que sugiere que el flujo proviene de una región no mayor que la distancia que la luz recorre en un día. Tal distancia es solamente diez veces mayor que el diámetro de un agujero negro cuya masa multiplique por mil millones la solar. Para las energías en la banda de rayos X se han registrado grandes cambios de flujo en menos de una hora. En segundo lugar, un grupo de astrónomos ópticos dirigidos por Wallace L. W. Sargent y Peter J. Young, del Cal Tech, han investigado la luz procedente de las galaxias M87 y NGC 6251. Arguyen que la distribución y el espectro de la luz indican la presencia de una masa compacta que excede miles de millones la solar. En tercer lugar, un agujero negro representa el más profundo "pozo" gravitatorio posible. De aquí que un agujero negro sea el objeto mejor capacitado para convertir eficazmente la energía del gas que cae hacia su interior en energía radiante o cinética.



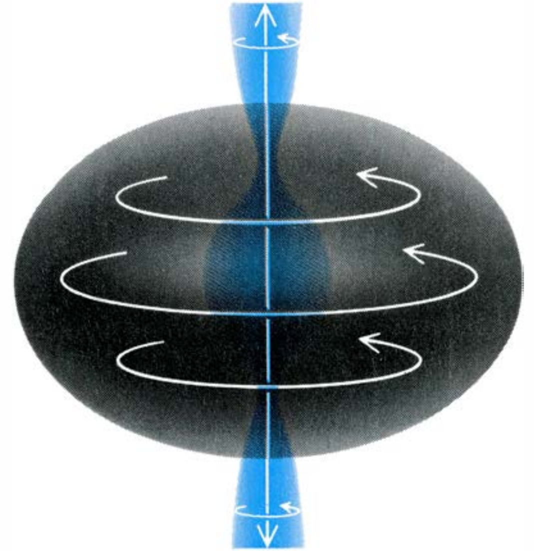
LA CIZALLADURA EN EL CHORRO tiene un efecto diferente sobre el campo magnético del mismo: hace la componente paralela dominante. La cizalladura, causada por falta de uniformidad en la velocidad del gas en el chorro (*flechas blancas*), alarga los bucles en el campo perpendicular. Las regiones alargadas pueden terminar por desprenderse de los bucles padres y cerrarse para convertirse en pequeños bucles.

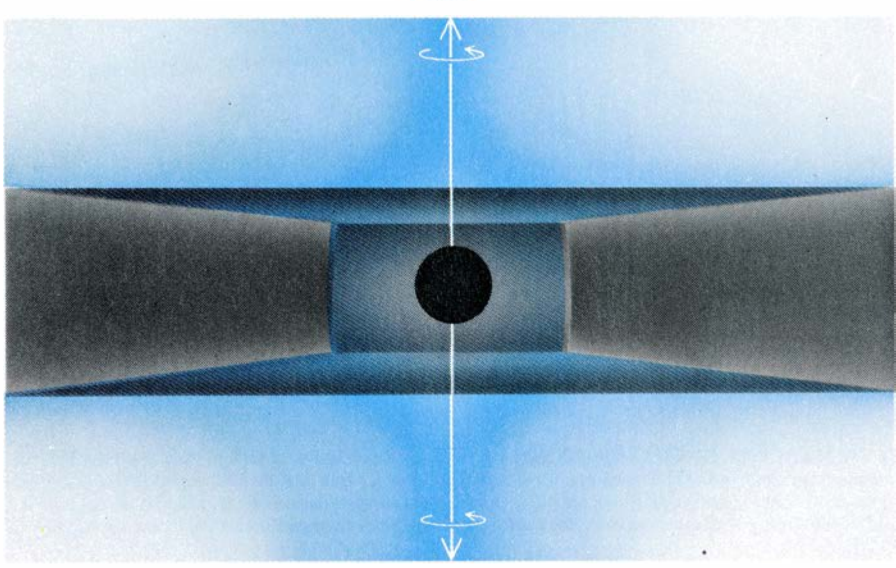
Si el agujero negro está en rotación, distorsionará el espacio circundante a través de un efecto que puede predecirse en el marco de la teoría general de la relatividad. La distorsión hará que el gas que rodea el agujero realice una precesión alrededor del eje de rotación del agujero. A diferente distancia, el gas tendrá velocidades de precesión distintas. A resultas de ello ocurrirán dos cosas. En primer lugar, el rozamiento viscoso entre anillos de gas vecinos hará que el gas se compacte hasta formar un disco en el plano ecuatorial del agujero. En segundo lugar, la misma resistencia viscosa que produjo el disco hará que la materia se dirija lentamente hacia la parte interior de éste. El disco no sólo aportará masa a los chorros, sino que les suministrará, además, potencia mediante la liberación de energía gravitatoria. Presumiblemente, los chorros saldrán expulsados en direcciones paralelas al eje de rotación del agujero negro; direcciones que son bastante estables, habida cuenta de la gran dosis de momento cinético que posee el agujero negro en rotación.

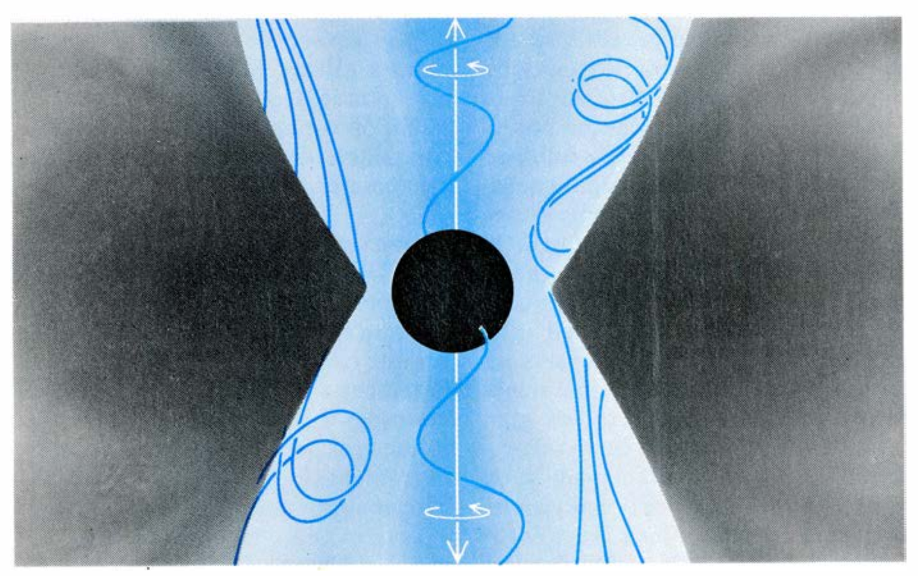
La estructura del disco de acumulación estará regida por la velocidad a la que la resistencia viscosa le suministra el gas. Si la velocidad de suministro es grande, la presión de radiación producida en la parte más interna del disco hinchará ésta. El agujero negro quedará entonces rodeado por un grueso toro de gas, o volumen de gas en forma de rosquilla. La superficie del toro próxima al eje de rotación del agujero definirá dos estrechos embudos en cada uno de los cuales puede colimarse un chorro. Los chorros escaparán a través de sus embudos impulsados, en gran parte, por la presión del campo de radiación en el interior del embudo y, en parte quizá también, por la energía mecánica que ha transportado hasta la superficie del embudo la convección del gas en su interior. Persiste, no obstante, una dificultad. Un disco mantenido por la presión de radiación supone una luminosidad óptica y ultravioleta mucho mayor que la generalmente observada en el centro de las galaxias asociadas a radiofuentes dobles. Pese a ello, hay razón para pensar que en SS433, el objeto emisor del chorro dentro de nuestra galaxia, la provisión de gas es muy grande. Un toro mantenido por la radiación continúa siendo una explicación viable de que broten chorros de esta fuente.

Una variante del toro mantenido por la radiación es la del toro mantenido por gas, que puede producirse a una tasa de suministro de gas pequeña. En este caso, la radiación generada en la zona más interna del disco escapa libremente y no puede inflarlo. Pero el rozamiento viscoso del disco es suficientemente grande para que el gas libere energía gravitatoria cayendo hacia el agujero negro más deprisa de lo que puede radiar energía. Esto significa que el gas se calienta a medida que cae. El calentamiento lo hace llenar un volumen tórico.

La potencia de los chorros no tiene necesariamente que proceder del gas que cae hacia dentro; puede extraerse del propio agujero negro. En particular, de un agujero negro en rotación se







HIPOTESIS SOBRE EL ORIGEN DE LOS CHORROS. Varios son los mecanismos postulados para la formación de esos haces. El mecanismo de boquilla (*arriba*) supone que una nube giratoria de gas rodea a una nube boyante de gas más caliente. Este escapa por el eje de rotación de la nube. Según el mecanismo de disco de acumulación (*centro*), el campo gravitatorio de un agujero negro de masa superior a lo normal confina el gas en un disco. Los chorros, alimentados por la presión resultante de la radiación electromagnética del gas, son lanzados a lo largo del eje de rotación del agujero. El mecanismo de toro de acumulación (*abajo*) supone que la presión de radiación (o la presión del gas caliente) hincha la parte interna del disco de acumulación alrededor de un agujero negro. Los chorros obtienen su energía de esta presión, de la energía convectiva en el gas inflado o del agujero negro en rotación, que suministra energía a los chorros arrastrando a las líneas del campo magnético a través de ellos. (Ilustración de Ian Worpole.)

puede extraer una cantidad significativa de energía mecánica. Supongamos que el gas que fluye hacia el agujero posee un campo magnético. Con el tiempo, algunas de las líneas del campo penetran en la superficie del agujero. Las líneas quedan débilmente ligadas al agujero; a medida que éste gira, arrastra a las líneas del campo consigo. Estas líneas pueden realizar así un trabajo sobre la materia próxima al agujero negro. De hecho, las líneas de campo ligadas al agujero son bastante parecidas a las cuerdas arrolladas alrededor de un torno. En cuanto al agujero negro en sí, actúa como un volante gigante cuya velocidad de rotación va reduciéndose gradualmente.

Todavía no hemos abordado el punto referente a cómo se suministra combustible a los chorros cósmicos. Fundamentalmente, las radiofuentes más potentes requieren varias masas estelares de gas cada año, a menos que la potencia de la fuente provenga en parte de un agujero negro en rotación. Los procesos normales de evolución estelar en la galaxia que rodea al núcleo de la radiofuente cumplen en cierto grado este requisito. Las mejores estimas disponibles sugieren, sin embargo, que se quedan cortos en al menos un factor de 10. Además, la violenta actividad del núcleo probablemente expulsará gas. Una hipótesis diferente es la que afirma que los potentes campos gravitatorios del centro de una galaxia crearán un denso cúmulo de estrellas. Las estrellas pueden presentar allí una distribución 10 millones de veces más densa que las estrellas ubicadas en la vecindad del Sol. Tal acumulación significa una probabilidad mucho mayor de que las estrellas choquen. Las estrellas pueden destruirse, además, por fuerzas de marea si se acercan demasiado a un agujero negro de gran masa. Pese a ello, ciertos cálculos detallados han puesto de relieve la dificultad de que llegue a formarse y persista el cúmulo de estrellas necesario.

Una tercera hipótesis recaba su apoyo observacional de las galaxias con centro doble o incluso múltiple. Se trata de un proceso al que Jeremiah Ostriker, de la Universidad de Princeton, ha llamado canibalismo. El centro de una galaxia "caníbal" consume el centro de una pequeña galaxia compañera: la galaxia "misionera". Las porciones externas de la misionera las fraccionan rápidamente las fuerzas de marea. Más difícil resulta digerir las partes internas. Así, caen éstas hacia el centro de la caníbal. La idea de que el canibalismo ali-

menta a los chorros es atractiva porque las radiogalaxias abundan más a grandes distancias. Y menudeaban más en el pasado que ahora; en particular, en los tiempos en que las galaxias estaban próximas entre sí y tenían la máxima probabilidad de interaccionar. Una galaxia poseía una probabilidad de ser una radiofuente 1000 veces mayor cuando el universo tenía la cuarta parte de edad que ahora. Se sabe, además, que, entre las galaxias próximas, las que se encuentran donde la densidad de galaxias es elevada tienen mayor probabilidad de ser radiofuentes que las que se hallan aisladas.

Admitiendo que las radiofuentes se alimentaran por canibalismo, la explicación de los chorros de simetría inversa surgiría espontánea. Supongamos que tanto la galaxia misionera como la caníbal albergan un agujero negro central de masa superior a lo normal; el agujero negro misionero quedará establecido con el tiempo en una órbita binaria alrededor del caníbal. Supongamos también que el agujero negro misionero mantiene chorros colimados a lo largo de su eje de rotación. Normalmente, el agujero apuntará su eje de rotación a lo largo de una dirección fija en el espacio. Pero ahora, el campo gravitatorio del agujero negro compañero perturbará el espacio en su vecindad. A raíz de lo cual, el observador distante hallará que el eje de rotación tiene movimiento de precesión. Y si el eje lo tiene, también los chorros.

Gracias principalmente a la construcción de los grandes interferómetros interconectados y al desarrollo simultáneo de la interferometría de muy larga base, los radioastrónomos han dado grandes pasos hacia la revelación de la estructura de las radiofuentes dobles, chorros cósmicos incluidos. En lo que se refiere a la interpretación, sin embargo, seguimos envueltos en la ignorancia. Nos falta una idea clara de la composición de los chorros, de la velocidad a que se mueven, de la manera como están confinados y de por qué son tan estables. Lo más importante de todo: no sabemos qué es lo que los produce. Se manejan varias hipótesis. Algunas ingeniosas, otras improbables; pero las hay muy razonables.

La omnipresencia de los chorros cósmicos sugiere que todo lo que un núcleo galáctico necesita para expulsar sus escapes en direcciones opuestas es la materia prima de un profundo pozo gravitatorio, una provisión de gas y cierta velocidad de rotación. Pese a ello, los chorros son heterogéneos. Algunos parecen rápidos, otros lentos. Algunos se muestran potentes, otros débiles. Los hay largos y cortos. Tal vez los teóricos han venido insistiendo demasiado en hallar un mecanismo que dé cuenta de todo ello. Un conjunto distinto de procesos podría hallar sus contrapartidas en la naturaleza y conducir siempre a chorros cósmicos.

¿Cuáles son las perspectivas de progreso en el estudio de los chorros cósmicos? En cuanto a la parte de observación, el futuro se presenta prometedor. La comparación de los chorros extragalácticos con los chorros de nuestra galaxia proporcionará razones para discenir qué puntos relativos a la producción y propagación de los chorros no dependen de la escala. Además de SS433, la fuente de rayos X Escorpión X-1 (sólo a 600 años-luz de distancia) muestra indicaciones claras de doble estructura. Además, cierto número de nubes de gas interestelar y polvo han ofrecido desesperantes indicios de que el origen de los chorros tendría que ver con la formación de las estrellas.

Hasta ahora, la Gran Distribución ha cartografiado sólo una pequeña fracción de las radiofuentes que le son accesibles; muy pocas de entre ellas, a la máxima sensibilidad y resolución disponibles. Además, la técnica de interferometría de muy larga base ha llegado a su edad adulta. El próximo paso será la construcción de grandes conjuntos de radiotelescopios dedicados a VLBI. En longitudes de onda ópticas, cabe esperar mucho del lanzamiento del Telescopio Espacial en 1985. El aparato, que tendrá más de 20 veces el poder de resolución de los telescopios de la Tierra, nos adentrará en la naturaleza de los quasars y descubrirá varios chorros más a longitudes de onda ópticas.

En cuanto a la parte teórica, queda un largo trecho por andar. Con los ordenadores se pueden realizar cálculos detallados del flujo de gas alrededor de agujeros negros de gran masa. Tales estudios tropiezan, no obstante, con nuestra ignorancia sobre las propiedades microscópicas del plasma bajo condiciones cósmicas. El enfoque experimental puede resultar muy fructífero. En particular, los físicos aerodinámicos realizan habitualmente ensayos con chorros de gas. La velocidad de los haces con que trabajan apenas superan pocas veces la del sonido en ellos. Quizá se precisen velocidades mayores. El que las extrañas y bellas formas que detectan en el cielo los radiotelescopios se lleguen a reproducir en un túnel aerodinámico constituye una emocionante posibilidad.

El calendario gregoriano

Gregorio XIII estableció el calendario moderno hace 400 años con el fin de corregir el desfase acumulativo del calendario juliano y poder seguir celebrando la Pascua en primavera

Gordon Moyer

En 1582 Gregorio XIII introdujo el calendario que lleva su nombre; es el sistema cronológico de uso civil en todo el mundo. Antes de adoptar el calendario gregoriano, la civilización occidental dependía de un sistema, denominado calendario juliano, establecido por Julio César. Este sistema se utilizó durante más de 16 siglos, pese a una divergencia acumulativa entre la duración media del año civil, 365.25 días, y la del año trópico, que es el intervalo medio de tiempo entre tránsitos sucesivos del sol aparente por el equinoccio de primavera. En 1582, el error del sistema juliano era del orden de 11 días, aproximadamente. Este efecto constituía un motivo de preocupación del Pontífice: si se hubiera seguido utilizando el calendario juliano, la Pascua hubiera acabado por celebrarse en verano.

El Papa Gregorio reunió un equipo eminente de astrónomos, matemáticos y clérigos para llevar a cabo la reforma del calendario juliano. La comisión tuvo que hacer frente a un problema fundamental inherente a todos los calendarios civiles: por razones obvias, un calendario diseñado para el uso ordinario debe constar de un número entero de días; no puede dejar pendiente una fracción de día al final del año. Uno de los miembros más significados de la comisión, el astrónomo jesuita Christoph Clavius, planteó esta idea de modo sucinto: "Annum civilem necessario constare ex diebus integris" ("Los años civiles deben constar necesariamente de días íntegros"). Este requisito básico es la raíz de todas las dificultades para la construcción de un calendario exacto. También es la razón fundamental por la cual resulta imposible crear un calendario perfecto; a saber: aquel que nunca pueda errar ni siquiera por un solo día.

En principio, el calendario gregoriano consiste en una mera versión ligeramente modificada del calendario juliano. La comisión decretó la eliminación de diez días del año 1582 con el fin de hacer coincidir de nuevo la fecha del equinoccio de primavera con el 21 de marzo; a lo largo de muchos siglos, la fecha había retrogradado al 11 de marzo. Para corregir el desfase de la fecha del equinoccio se adoptó un plan en virtud del cual en tres de cada cuatro años centenarios –1700, 1800 y 1900 por ejemplo– se omitiría el día intercalar que debía añadirse en aplicación del sistema juliano. Estas revisiones, promulgadas en la bula papal del 24 de febrero de 1582, dieron lugar a una amplia polémica, promovieron un furioso debate entre los eruditos y motivaron el que el hombre de la calle se preguntara si las aves sabrían, entonces, en qué momento debían volar hacia el sur para invernar. No hay más que hojear la gran *Bibliographie Générale de l'Astronomie*, publicada en 1887 por Jean-Charles Houzeau y Albert-Benoit Lancaster, para hacerse una idea del número enorme de tratados que se escribieron en favor y en contra de la reforma.

La controversia abarcaba por igual el campo religioso y el académico. Nos encontramos en la época de la Reforma: los países protestantes rechazaron el nuevo calendario al que denunciaron como una argucia del Papa para lograr que el rebaño rebelde volviera a caer bajo la jurisdicción de Roma. Esta acusación no carecía por completo de fundamento. Gregorio XIII era un promotor vigoroso, o mejor despiadado, de la Contrarreforma. Ayudó a Felipe II de España a hostigar a los protestantes holandeses y aceptó jubiloso la cabeza del caudillo de los hugonotes franceses tras la infame matanza del Día de San Bartolomé, baño de sangre que el Papa celebró como una victoria católica, ordenando que se acuñara una medalla para conmemorarlo. Probablemente Gregorio consideró que era un momento oportuno para imponer una reforma calendárica al mundo cristiano y la llevó a cabo amenazando con la excomunión a los que rehusaran aceptarla.

La oposición al calendario no surgió exclusivamente como consecuencia de la contienda religiosa de la época. Muchos hombres cultos reconocían la necesidad de una reforma del calendario (el error acumulativo del calendario juliano había sido denunciado durante siglos), pero no estaban persuadidos de que el sistema gregoriano constituyera un avance significativo con respecto al viejo sistema. De hecho, el eminente matemático François Viète, al que con frecuencia se considera el padre del álgebra moderna, condenó la reforma como una corrupción del calendario juliano.

Algunos científicos sobresalientes del siglo XVI, entre los que se encontraba el propio Viète, afirmaban que el calendario gregoriano carecía de una adecuada base astronómica. De esta opinión participaban dos de los críticos más duros de la reforma, Michael Maestlin y Joseph Justus Scaliger. Maestlin, astrónomo, fue uno de los primeros que abrazó abiertamente la teoría copernicana y debe su fama al hecho de haber sido profesor de Johannes Kepler en Tübingen. Scaliger, otro académico de fama, fue un extraordinario erudito que dominaba una docena de lenguas: buen conocedor de las lenguas clásicas, historiador, filólogo y cronologista. Sus colegas lo consideraban un "mar de ciencias" y un "pozo sin fondo de erudición". Se convirtió en el archienemigo de Clavius, el principal defensor de la reforma gregoriana.

En 1595 Scaliger publicó un comentario al *Canon Paschalis* de Hipólito, una obra del siglo IV sobre el cómputo de la fecha de la Pascua. Añadió al comentario un apéndice consistente en



DETALLE DE UN CUADRO que pone de relieve la causa de la reforma del calendario juliano. Un miembro de la comisión nombrada por Gregorio XIII en 1576 para estudiar la necesidad de una reforma aparece señalando un gráfico que muestra el error de diez días que, según habían establecido los astrónomos, se había acumulado en el calendario juliano desde la época del primer Concilio Ecuménico de Nicea en el año 325. El arco superior del diagrama representa un segmento del año calendárico dividido en días, mientras que el arco inferior muestra una sección correspondiente del año trópico (el año medido por pasos sucesivos del sol por el equinoccio vernal). Los signos zodiacales de Libra y Scorpio indican el mes de octubre, del que se tomaron los diez días sobrantes. Fn el cuadro entero aparece el Papa sentado en un trono a la izquierda. La obra pertenece a un pintor anónimo, cuyo encargo recibió en 1582, el año de la reforma, de Scipio Turamini, un miembro del concejo de Siena. La pintura se conserva en el Archivo de la Ciudad de Siena.

una crítica mordaz del calendario gregoriano y que era dos veces más extenso que la obra principal. Scaliger defendía su propio plan de reforma que hubiera resultado en un calendario algo más preciso, pero también un tanto más complicado.

Clavius respondió rápidamente a las afirmaciones de Scaliger. En el mismo año publicó una réplica punzante, *Iosephi Scaligeri elenchus, et castigatio calendarii Gregoriani*. En ella las críticas de Scaliger, tomadas de su obra sobre el *Canon Paschalis*, y las respuestas de Clavius aparecen en párrafos alternativos. Las observaciones de Clavius son frecuentemente divertidas por su vehemencia pero resultan rara vez más ofensivas que las invectivas de Scaliger. Fustigó al bávaro Clavius llamándole "barrigón alemán" y "bestia".

Scaliger hizo objeciones tanto a la parte civil como a la eclesiástica del calendario. Afirmó que las tablas utilizadas para computar la fecha de la Pascua, las llamadas tablas de epactas, eran erróneas y consideró que las reglas reformadas para determinar los años bisiestos no eran suficientes para mantener la fecha del equinoccio de primavera fija en el 21 de marzo. En realidad, el calendario implica cierta fluctuación en la fecha del equinoccio que llega a alcanzar unos pocos días. A pesar de ello, el calendario logra mantener la fecha en el día 21 de marzo, o en sus proximidades, y seguirá manteniéndolo durante muchos siglos.

Como Scaliger, tampoco Maestlin aprobó el sistema gregoriano para la determinación de la Pascua. El Concilio de Nicea del año 325 decretó que la Pascua debía ser celebrada por todos los cristianos en el mismo día. La Iglesia determinó que la fecha debía ser el primer domingo posterior al día 14 del mes lunar (aproximadamente la luna llena) que coincide con o sigue al equinoccio de primavera, el cual, en tiempos del concilio, se consideraba fijo en el 21 de marzo. Esta regla engorrosa se aplica todavía hoy para establecer la fecha de la Pascua y, por consiguiente, las de todas las fiestas móviles. Computar la fecha de la Pascua es, sin duda, la operación más compleja a realizar con el calendario. Ni siquiera Carl Friedrich Gauss, que estudió el problema en su juventud, logró desarrollar un algoritmo completo capaz de derivar la fecha.

En un año cualquiera, la Pascua puede caer en cualquier domingo comprendido entre el 22 de marzo y el 25 de abril. Aunque parezca increíble, hace falta que transcurran 5.750.000 años para que todas las fechas de la Pascua se repitan en el mismo orden. La tabla de epactas se utiliza para determinar la edad de la luna el 1 de enero de cada año en un ciclo de 7000 años. Ya que el intervalo medio entre una luna nueva y la siguiente, o sea una lunación, es algo más de 29,53 días, el número máximo de epactas es 30 (29,53 redondeado al número entero inmediatamente superior).

Si conocemos la edad de la luna el 1 de enero, resulta muy sencillo determinar las fechas de todas las lunas nuevas y lunas llenas del año. Será fácil, entonces, descubrir cuál es el primer día 14 del mes lunar a partir del 21 de marzo. La Pascua será el primer domingo que siga a este día 14. El ciclo de epactas, no obstante, es sólo una aproximación a las lunas verdaderas o astronómicas; una epacta puede divergir hasta en tres días con respecto a la fase lunar real, por más que la desviación habitual suele ser de uno o dos días. La primera ventaja de un ciclo lunar artificial es su simplicidad. El movimiento verdadero de la luna resulta demasiado complejo para reducirlo a una tabla sencilla y adecuada, susceptible de ser entendida con facilidad por clérigos peritos en el cómputo, la ciencia que, durante la Edad Media y el Renacimiento, se ocupaba de calcular la fecha de la Pascua.

Clavius se mostraba satisfecho con aproximaciones suficientemente ajustadas. Urgió repetidamente la necesidad de un calendario sencillo, observando que resultaba absurdo pretender que todo el mundo dominara la astronomía. Pero Maestlin defendió el uso de cálculos astronómicos exactos para derivar la fecha de la luna llena pascual. Kepler adoptó una postura contraria a la de su antiguo maestro y apoyó el método de Clavius para seguir el curso de la luna, que era razonablemente preciso, argumentando que "la Pascua es una festividad y no un planeta".

Debido sobre todo a su incansable defensa del calendario, Clavius se ganó un puesto de honor en la historia de la ciencia, hasta el punto de que el mayor cráter de la luna lleva su nombre. En conjunto escribió cinco defensas del calendario. La primera, publicada en 1588, fue una respuesta a las críticas de Maestlin. La obra definitiva de Clavius sobre el calendario es, no obstante, un tomo de 800 páginas titulado *Romani calendarii a Gregorio XIII P.M. restituti explicatio*, publicado por



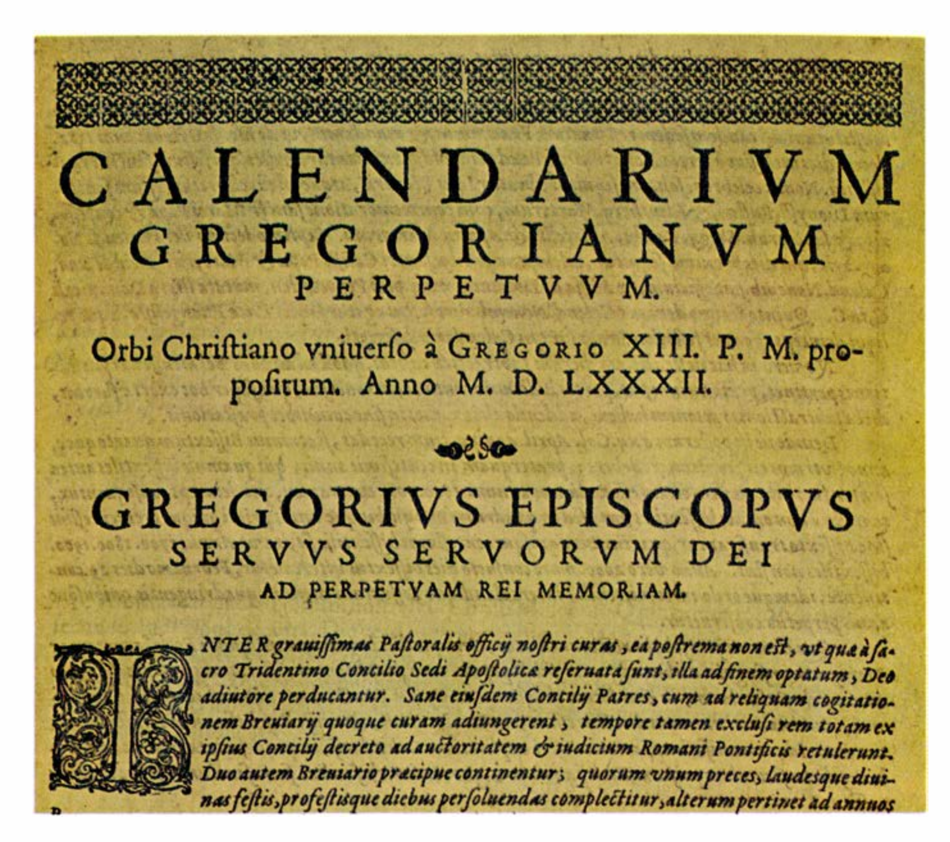
ESTATUA DE GREGORIO XIII, situada sobre la entrada del Palazzo Pubblico en Bolonia. El portal fue diseñado por Galeazzo Alessi, arquitecto del siglo XVI, para conmemorar la elección papal ya que Gregorio había nacido en esa capital italiana. Accedió al solio pontificio el día 13 de mayo de 1572 y diez años más tarde estableció la reforma de un calendario juliano que llevaba ya 1600 años de existencia.

vez primera en 1603. Muchos de sus argumentos en contra de Maestlin se repiten, a menudo literalmente, en este tratado.

Clemente VIII, que accedió al papado en 1592, había encargado a Clavius que explicara y justificara el nuevo calendario ante el mundo cristiano. La *Explicatio* es el voluminoso resultado de este encargo. Se trata de una obra técnica de gran envergadura. El erudito de Edimburgo Alexander Philip, autor de uno de los escasos textos importantes escritos en inglés sobre el calendario, sostuvo que Clavius, "como un pulpo", inundó el tema con "el océano de tinta del que lo rodeó". Esta crítica tiene alguna base: Clavius es a menudo retorcido y repetitivo, pero rara vez resulta inaccesible; y otros eruditos han señalado que siempre es legible. De hecho, resultó uno de los autores científicos más leídos del Renacimiento tardío.

Todos los tratados de Clavius sobre el calendario fueron reeditados y reunidos en el tomo V de sus *Opera mathematica*, que fue publicado en 1612, el año de su muerte. El tomo V incluye también una reimpresión del *Kalendarium Gregorianum perpetuum*, la publicación oficial sobre la reforma que la Iglesia editó en 1582 y que había sido escrita por la comisión constituida por el Papa Gregorio. Se reimprime también la bula papal *Inter gravissimas*... ("Con la máxima ansiedad...") del 24 de febrero de 1582, en la que se anunciaba a todos los príncipes cristianos que el nuevo sistema empezaría oficialmente el 15 de octubre de aquel año. Se prescribía la supresión de diez días del calendario: "Con el fin de restaurar, por tanto, el equinoccio vernal en su lugar primitivo, que los Padres del Concilio de Nicea hicieron coincidir con el *XII Kalendas Aprilis* [21 de marzo], prescribimos y mandamos en relación con el mes de octubre del año 1582 que se supriman los 10 días comprendidos entre el *III Nonas* [5 de octubre] y el día que precede a los *Idus* [14 de octubre], ambos inclusive". Clavius y los restantes miembros de la comisión evitaron una discontinuidad en los días de la semana en esta "expulsión" de días: el jueves 4 de octubre de 1582, la última fecha del calendario juliano, fue seguido inmediatamente por el viernes 15 de octubre.

En su defensa frente a Maestlin, Clavius explicó que no existía ningún misterio en la elección del mes: octubre, sencillamente, era el mes con menor número de fiestas religiosas, razón por la cual la omisión de unos días en él causaba pocos problemas a la Iglesia. Lo describió asimismo como el mes que menos perturbaciones ocasionaba en la esfera comercial.


CALENDARIVM GREGORIANVM PERPETVVM.

Orbi Christiano vniuerso à GREGORIO XIII. P. M. propositum. Anno M. D. LXXXII.

GREGORIVS EPISCOPVS SERVVS SERVORVM DEI AD PERPETVAM REI MEMORIAM.

INTER grauissimas Pastoralis officij nostri curas, ea postrema non est, vt quæ à sacro Tridentino Concilio Sedi Apostolicæ reseruata sunt, illa ad finem optatum, Deo adiutore perducantur. Sane eiusdem Concilij Patres, cum ad reliquam cogitationem Breuiarij quoque curam adiungerent, tempore tamen exclusi rem totam ex ipsius Concilij decreto ad auctoritatem & iudicium Romani Pontificis retulerunt. Duo autem Breuiario præcipue continentur; quorum vnum preces, laudesque diuinas festis, profestisque diebus persoluendas complectitur, alterum pertinet ad annuos

BULA PAPAL del 24 de febrero de 1582 por la que se ordenaba a los cristianos de Europa que adoptaran el calendario gregoriano el 15 de octubre de 1582 (o en 1583). La bula fue reimpresa en el tomo V de las *Opera mathematica* de Christoph Clavius, jesuita astrónomo que fue miembro destacado de la comisión papal que estudió la reforma del calendario y que se convirtió más tarde en el principal defensor de la reforma. Esta fotografía muestra la primera página de la bula tal como aparece en el escrito de Clavius.

La eliminación de 10 días no constituía una necesidad astronómica. La fecha del equinoccio de primavera podría haber sido estabilizada en el 11 de marzo. El problema real era impedir que la fecha escogida, fuera cual fuera, continuara separándose de la fecha astronómica. Por consiguiente, ¿cómo logró el plan, puesto en marcha por el Papa Gregorio, mantener la fecha alrededor del 21 de marzo? Esta es la pregunta técnica esencial que uno debe plantearse al considerar la reforma. Aunque la mayor preocupación de la comisión gregoriana había sido restablecer la fecha de la Pascua, no podía alcanzarse este objetivo hasta que se hubiera estabilizado la fecha del equinoccio.

Verosímilmente, no se hubiera llevado a cabo ninguna reforma en 1582 si no hubiera entrado en juego la celebración de la Pascua. El error de unos 11 días que se había acumulado, debido al calendario juliano, entre la época del Concilio de Nicea y 1582, constituía aún una discrepancia de menor cuantía que no se había revelado todavía como una diferencia apreciable entre los días del calendario y las estaciones. De hecho, si el calendario juliano hubiera seguido en uso hasta el momento presente, su error sería sólo de unas dos semanas y todavía no resultaría apreciable para los habitantes del hemisferio norte como un desplazamiento de las fechas del calendario desde la primavera hacia el verano. El error, no obstante, resultaba significativo en relación a la Pascua ya que su observancia depende de una fecha fija del equinoccio de primavera.

La regresión continua de la fecha del equinoccio era, casi enteramente, una consecuencia de la discrepancia inicial entre la longitud del año juliano y la del año trópico. Por consejo del erudito greco-egipcio Sosígenes, Julio César, al establecer el calendario juliano en el año 45 a. de C., instauró un plan en virtud del cual cada año común constaba de 365 días y cada año divisible por cuatro con resto cero constaba de 366. El día intercalar se añadía, en un principio, en la víspera del 25 de febrero, denominada *ante diem sexto Kalendas Martius*, el sexto día antes de las calen-

das de marzo. Este día intercalar se denominaba, por consiguiente, *bissexto* [segundo sexto] *Kalendas Martius*, de donde deriva el término *bisiesto*, una palabra que todavía hoy se utiliza para designar el año y el día intercalar. En la reforma gregoriana el día intercalar se retrasó al último día de febrero.

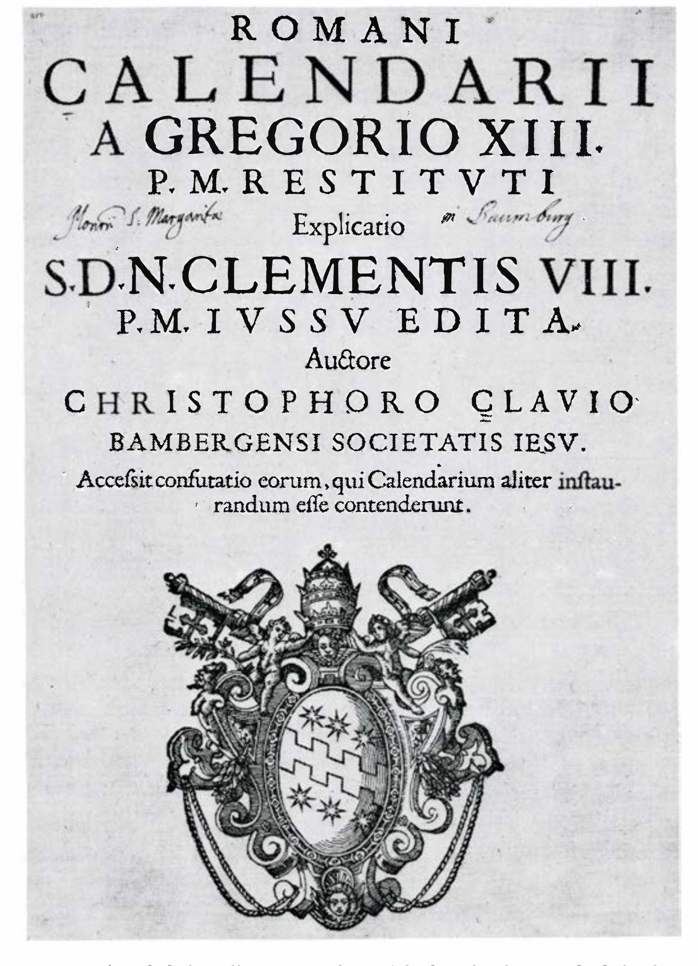
En el momento en que se creó el calendario juliano, la longitud del año trópico no era un parámetro bien establecido, al menos entre los astrónomos occidentales. El plan de Sosígenes para introducir un día suplementario cada cuatro años daba lugar a un año calendárico medio de 365,25 días. No obstante, en el año 45 a. de C., la longitud del año trópico era aproximadamente de 365,24232 días, o sea, aproximadamente 11 minutos y 4 segundos más corto que el año juliano. Esta diferencia, inicialmente muy pequeña, se fue acumulando hasta que, con el tiempo, no fue del orden de minutos sino de días. El error empezó entonces a revelarse como una regresión gradual de la fecha de los equinoccios y de los solsticios.

La acumulación del error en el calendario juliano se vio acelerada por la disminución gradual de la longitud del año trópico, un fenómeno sistemáticamente ignorado por los eruditos –incluso por los astrónomos– que trataban de los aspectos técnicos del calendario. Esta disminución tiene lugar en virtud de una progresión geométrica que debe tenerse en cuenta al analizar la precisión de un calendario solar. Si la diferencia inicial entre el calendario juliano y el sol medio, de 11 minutos y 4 segundos, se hubiera mantenido (si el calendario hubiera ganado este intervalo de tiempo cada año con respecto al sol), el error hubiera sido de un día solar medio en un período de tiempo ligeramente superior a 130 años. El desfase se hubiera mostrado como un retraso de la fecha del equinoccio de un día en relación a su posición 130 años antes. Pero, en realidad, la regresión de las fechas de los equinoccios y de los solsticios había sido más rápida. En la época en que tuvo lugar la corrección gregoriana el calendario erraba con respecto al sol medio a un ritmo de un día cada 128,5 años, aproximadamente.

La disminución de la longitud del año trópico resulta progresivamente más significativa a medida que transcurre el tiempo. Es una de las circunstancias que impide la creación de un calendario perfecto. La razón principal, no obstante, tal como indicó Clavius, es que un calendario diseñado para el uso ordinario debe proporcionar un número íntegro de días, mientras que el año trópico tiene una componente fraccionaria compleja. Todo ello se reduce a una cuestión de aritmética trivial: un valor fraccionario no puede nunca expresarse como un entero. Un calendario solar puede sólo aproximarse a la longitud del año trópico; la parte fraccionaria se toma en cuenta al utilizar la intercalación, esto es, mediante la inserción periódica de días intercalares o bisiestos en el calendario. Cuanto más se aproxime la intercalación a la parte fraccionaria del año, tanto más logrará el calendario mantener el equinoccio vernal en la misma fecha durante un largo período de tiempo.

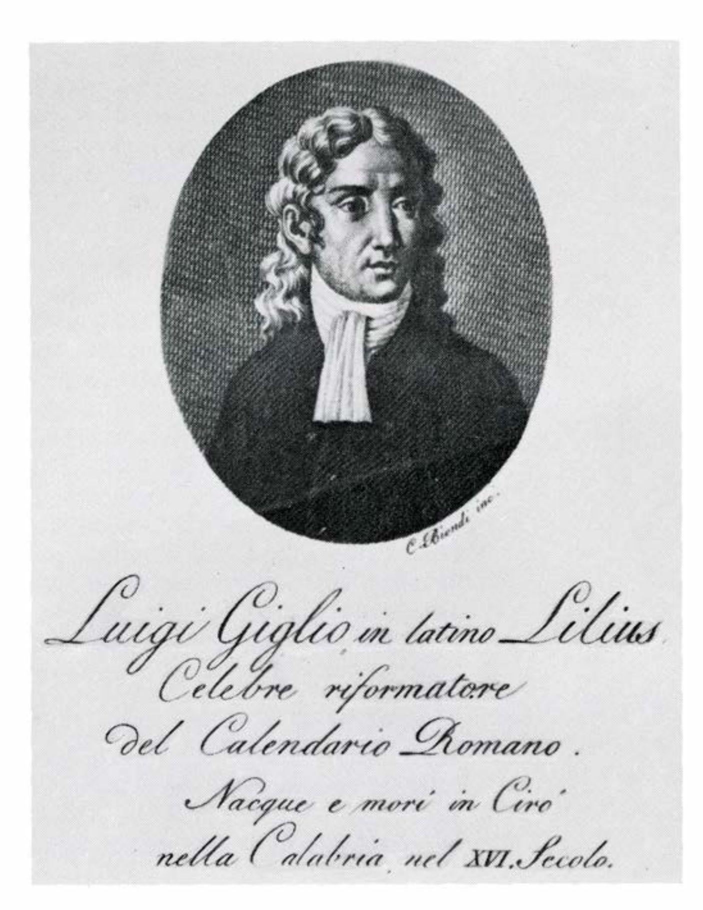
El plan de reforma adoptado por el Papa Gregorio engarza limpiamente con el sistema juliano y frena la





ROMANI
CALENDARII
A GREGORIO XIII.
P. M. RESTITVTI
Explicatio
S.D.N. CLEMENTIS VIII.
P. M. IVSSV EDITA.
Auctore
CHRISTOPHORO CLAVIO
BAMBERGENSI SOCIETATIS IESV.
Accessit confutatio eorum, qui Calendarium aliter instaurandum esse contenderunt.

CHRISTOPH CLAVIUS fue retratado en 1606, a la edad de 69 años, en un grabado (*izquierda*) debido a Francisco Villamoena. Fue un distinguido matemático que defendió con firmeza la reforma gregoriana. (En su honor, el mayor cráter de la luna lleva su nombre.) A la derecha, la portada de la obra definitiva de Clavius sobre el calendario, que le encargara Clemente VIII, quien accedió al papado en 1592; fue el quinto pontífice después de Gregorio.



PERITIS MATHEMATICIS.



VM in ſacro cōcilio Tridentino Breuiarij Miſſaliſque emendatio Romano Pontifici reſeruata eſſet, idque felicis recordationis Pius V. quanta maxima potuit diligentia ſuperioribus annis perficiendum curaſſet, atq; edidiſſet: nō tamen id opus uiſum eſt ſuis omnibus numeris abſolutum atque perfectum, niſi reſtitutio quoque anni & eccleſiaſtici Kalendarij accederet. In eam igitur curā dum Gregorius XIII Pont. Max. toto animo & cogitatione incumbit, allatus eſt illi liber ab Aloiſio Lilio cōſcriptus, qui neq; incommodam neque difficilem uiam ac rationem eius rei perficiendæ proponere uidebatur. Verum cū ea kalendarij emendatio multas ac magnas difficultates afferat, & iam diu a bonis uiris omnibus efflagitata, a doctiſsimis mathematicis ſæpe deliberata, & multū agitata, abſolui tamen adhuc, & ad exitum perduci minime potuerit, uiſum eſt prudentiſsimo Pontifici de ea re peritiſsimos quoſque huius ſcientiæ uiros conſulendos eſſe, ut res quę omnium communis eſt, communi etiam omnium conſilio perficiatur. Cogitarat itaque eum librum cunctis Chriſtianis Principibus mittere, ut ipſi adhibitis peritioribus mathematicis, illum aut ſua ſententia comprobarent, aut ſi quid deeſſe uideretur, id omne

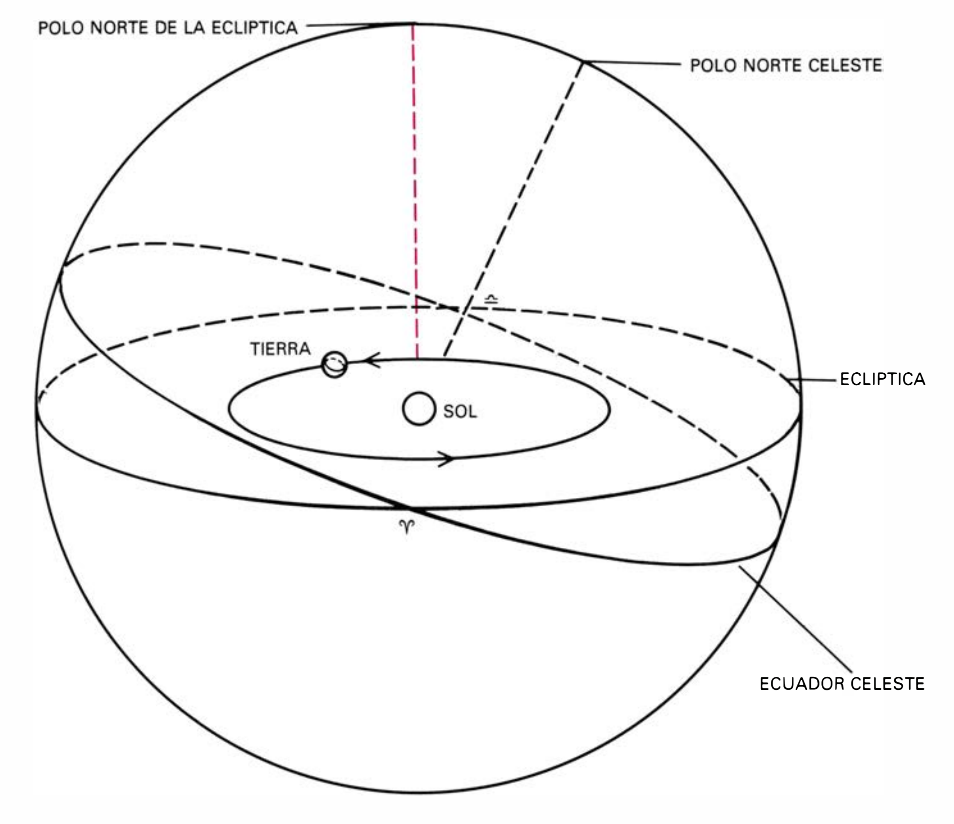
EL PADRE del calendario gregoriano fue Aloisius Lilius (forma latinizada de Luigi Giglio), que aparece retratado a la izquierda. A su muerte en 1576, su plan para la reforma del calendario fue rescatado del olvido por su hermano Antonio y sometido a la consideración del Papa Gregorio XIII. El plan fue adoptado por el Pontífice romano y por su comisión de asesores con una única modificación de escasa importancia debida a Christoph Clavius. A la derecha aparece la primera página del compendio de Lilius, donde expone su plan para el calendario. Esta obra fue descubierta por el autor de este artículo en la Biblioteca Nacional Central de Florencia, catalogada como de autor anónimo aunque en la página puede leerse: “Este libro fue escrito por Aloisio Lilio”. La *C* iluminada representa al Papa mientras preside una reunión de algunos de los miembros de la comisión para la reforma del calendario.

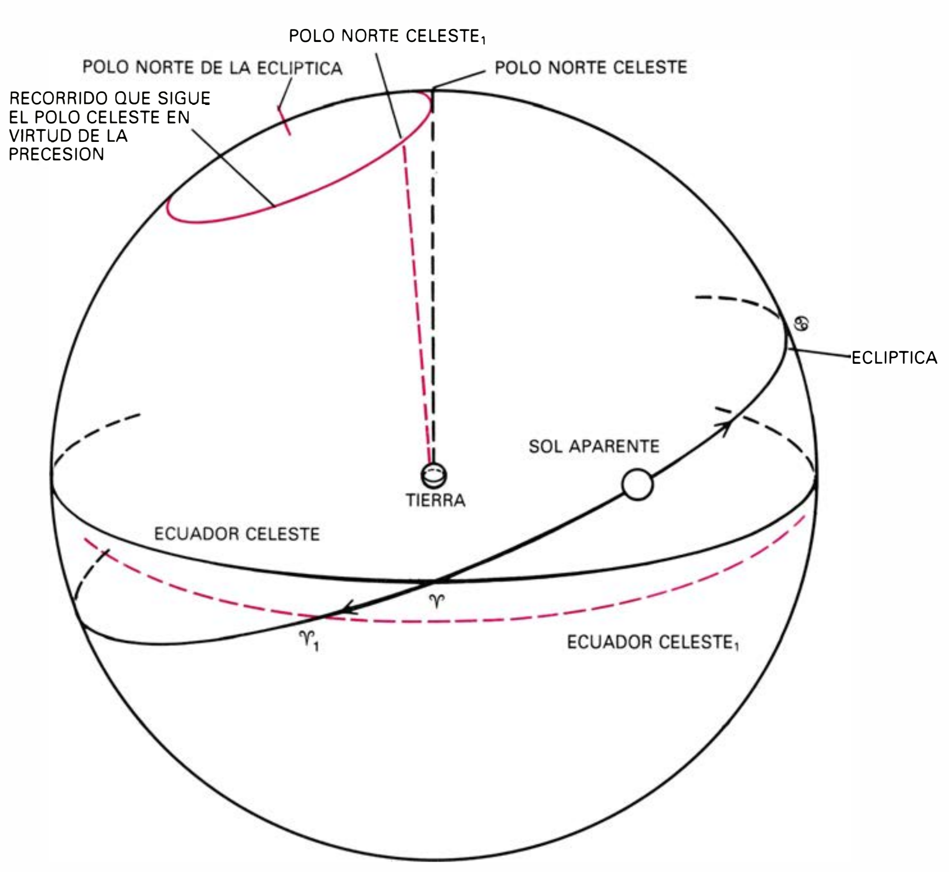
deriva de la fecha del equinoccio de modo tan eficaz que el calendario no perderá un solo día con respecto al sol en bastante más de 2000 años. El plan que constituye la base del calendario civil gregoriano no fue diseñado por Clavius ni por ningún otro miembro de la comisión, sino por un profesor de medicina de la Universidad de Perugia, el cual, desgraciadamente, no sobrevivió para ver su plan convertido en ley. Se llamaba Luigi Lilio (a menudo aparece escrito como Giglio), latinizado en Aloisius Lilius. Clavius le denominó el *primus auctor* del nuevo calendario. Antes de que fuera conocido universalmente como calendario gregoriano, se solía referir a él como calendario liliano. Lilio no sólo desarrolló la nueva intercalación para la comisión, sino que construyó también la tabla de epactas, modificada más tarde por Clavius para ajustarla mejor a la nueva regla de intercalación de años bisiestos.

Lilio recomendó la supresión de 10 días del calendario de una vez o a lo largo de un período de 40 años a partir de 1584, prescindiendo de los días intercalares que, normalmente, debían añadirse cada cuatro años durante ese intervalo. Lilio dejó a la comisión que optara entre estas dos alternativas. Según parece, fue Clavius quien decidió la supresión de los días de una vez en octubre.

Lilio trabajó unos 10 años en los detalles de su reforma. Consumió la mayor parte de sus esfuerzos en desarrollar un método para computar la Pascua que resultara conveniente y suficientemente exacto. Un problema de mucha menor envergadura era la construcción de un sistema de intercalación que se aproximara mejor a la longitud del año trópico. A su muerte, en 1576, su *Compedium novae rationis restituendi kalendarii* (*Compendio del nuevo plan para restablecer el calendario*), el documento históricamente más importante sobre el tema del calendario, fue presentado en manuscrito por su hermano Antonio Lilio al Papa Gregorio. Se le sometió, entonces, al examen de un grupo internacional de clérigos y eruditos distinguidos. La obra de Lilio fue alabada por su precisión y simplicidad y entre la multitud de proyectos que se habían presentado era claramente el favorito con vistas a su adopción.

(En una nota publicada en 1974 en el *Journal for the History of Astronomy*, Noel M. Swerdlow, de la Universidad de Chicago, afirmó que el tratado de Lilio parecía perdido. Añadió: “No es imposible, empero, que el *Compendium* sobreviva, aún por descubrir, en su forma manuscrita”. Recientemente, Swerdlow me escribió que Thomas Settle, del Instituto Politécnico de Nueva York, había logrado averiguar que una versión impresa podía encontrarse en los archivos de la Biblioteca Nacional Central de Florencia. Una copia se encuentra realmente en la mencionada biblioteca, donde aparece catalogada como obra de autor anónimo. No obstante, en la primera página de ese delgado volumen Lilio afirma explícitamente que él es su autor. Confirmada la existencia del *Compendium* en Florencia, lo busqué en otras ciudades italianas. Este libro raro, cuyo paradero, curiosamente, no mencionan nunca los eruditos, se encuentra también en los archivos de la Biblioteca Comunale degli Intronati en Siena y en la Biblioteca Vati-





POR AÑO TROPICO se entiende el intervalo medio de tiempo entre dos tránsitos consecutivos del sol aparente por el equinoccio vernal (*arriba*). El equinoccio vernal (♈) es la intersección de la eclíptica, el plano de la órbita terrestre, y el ecuador celeste (la proyección del ecuador terrestre sobre la esfera celeste). Para un observador terrestre el sol parece girar en torno a la Tierra (*abajo*). El sol en el equinoccio vernal se mueve sobre la eclíptica hacia el solsticio de verano (♋), mientras que el equinoccio vernal realiza su movimiento de precesión en el sentido opuesto. Al cabo de un año, el sol no ha llevado a cabo todavía una revolución completa en el momento en el que vuelve a pasar por el equinoccio vernal cuya posición se ha desplazado ($♈_1$). Debido a la fricción de las mareas, la Tierra disminuye gradualmente la velocidad de su rotación, lo cual da como resultado el correspondiente incremento en la precesión de los equinoccios. La consecuencia es que la longitud del año trópico está disminuyendo de manera gradual.

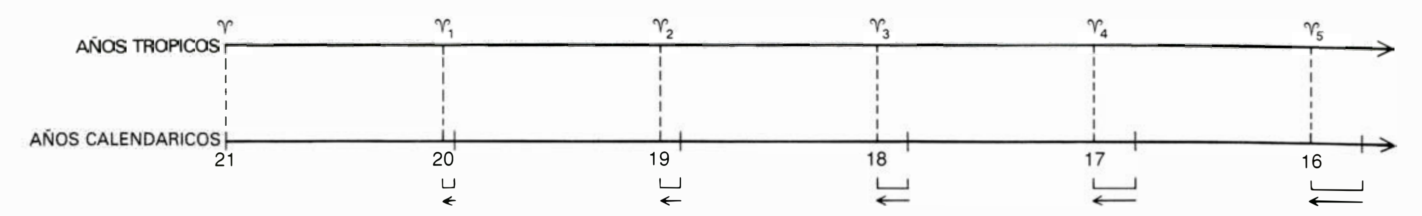
cana. El ejemplar de cada una de estas bibliotecas se encuentra encuadernado en una colección miscelánea de obras calendáricas publicadas por separado por contemporáneos de Lilio, el más conocido de los cuales es Alessandro Piccolomini.)

La intercalación que Lilio proponía en el *Compendium* era simple: la supresión de tres días intercalares en los años centenarios que no fueran exactamente divisibles por 400. La intercalación gregoriana sigue, por consiguiente, el sistema juliano con la única excepción de que años centenarios como 1700, 1800 y 1900, que habrían sido bisiestos en el calendario juliano, se convirtieron en años comunes (por omisión del día intercalar). El calendario gregoriano reduce el número de días intercalares a 97 en 400 años, mientras que 400 años julianos contenían 100 días intercalares.

El cambio introducido, aunque pequeño, es profundo. Logra que la longitud media del año calendárico se ajuste mucho mejor a la del año trópico, estableciendo un año medio calendárico de 365,2425 días. En 1582 la longitud del año trópico era aproximadamente de 365,24222 días, parámetro que difería del año gregoriano sólo por algo más de 24 segundos. Si la misma longitud del año no disminuyera gradualmente, el calendario basado en el plan establecido por Lilio mantendría la fecha del equinoccio de primavera en torno al 21 de marzo durante más de 3550 años.

El mismo Clavius estimó que el calendario cometería un error de un día en el año 5084. Si tomamos en cuenta la disminución de la longitud del año, algo totalmente desconocido por los astrónomos del siglo XVI, el calendario seguirá siendo exacto dentro del margen de un día solar medio durante un período de tiempo mucho más corto, unos 2417 años, o sea, aproximadamente hasta el año 4317. (Las extrapolaciones se basan en la determinación del año trópico en la época 1900 por el astrónomo americano Simon Newcomb; su cifra era de 365,24219879 días solares medios.)

Averiguar cómo llegó a su parámetro de 365,2425 continúa siendo un misterio. El *Compendium* no arroja, por desgracia, nueva luz sobre la cuestión. Los valores más fiables del año trópico conocidos por Lilio eran los que aparecían en las *Tablas Alfonsíes* de 1252, en el *De revolutionibus orbium coelestium* de 1543 y en las *Tablas Pruténicas* de 1551 de Copérnico. Cada una



LA RETROGRADACION en la fecha del equinoccio vernal ocurre cuando la duración media del año calendárico es más larga que la del año trópico. Se representan aquí varios equinoccios vernales. El primero cae el 21 de marzo pero, a medida que transcurre el tiempo, la discrepancia entre el año calendárico y el año trópico comporta la retrogradación que muestran los corchetes. Para simplificar, despreciamos la disminución gradual del año trópico.

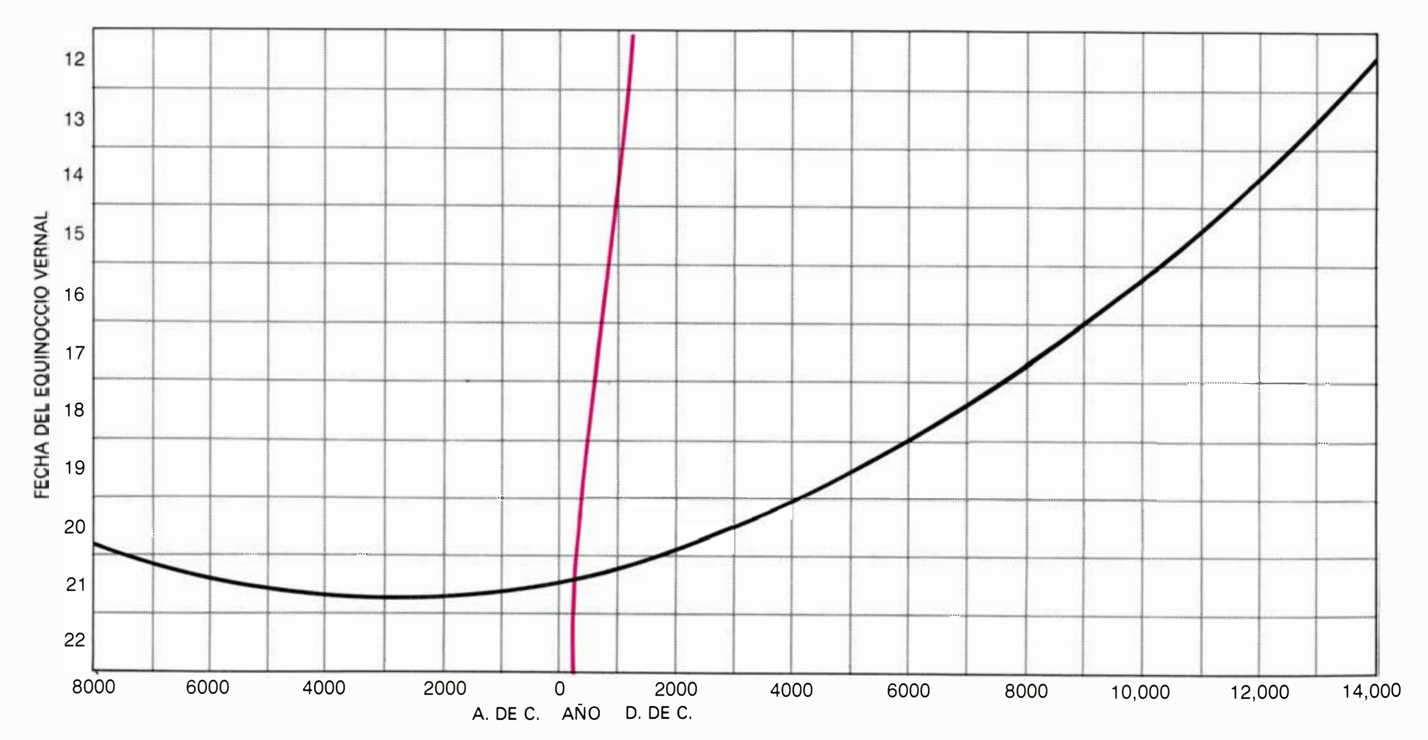
de estas fuentes da un valor de unos 365 días 5 horas 49 minutos y 16 segundos; difieren unas de otras en menos de un segundo, a pesar de lo cual son aproximadamente 4 segundos más largos que el año gregoriano. Ninguno de ellos da como resultado una intercalación de 97 días en 400 años. Swerdlow observó, no obstante, que los tres parámetros expresados en notación sexagesimal –la forma en la que se obtendrían a partir de una tabla del movimiento medio del sol– son idénticos hasta el segundo orden fraccionario sexagesimal. La manera de expresar la fracción sexagesimal es $365;14,33^{d}$ (el segundo orden fraccionario sexagesimal es 33). Si convertimos la fracción sexagesimal en una fracción común obtenemos $365^{97}/_{400}$, esto es, 365,2425 días. Es una teoría tan buena como cualquier otra para explicar el origen de la longitud del año gregoriano.

Lilio no fue el primero en proponer una intercalación de 97 días en 400 años. El mismo plan lo avanzó, en 1560, Petrus Pitatus, astrónomo veronés. No sabemos si Lilio tomó la idea de Pitatus, pero la intercalación es suficientemente sencilla como para que pueda suponerse una deducción independiente.

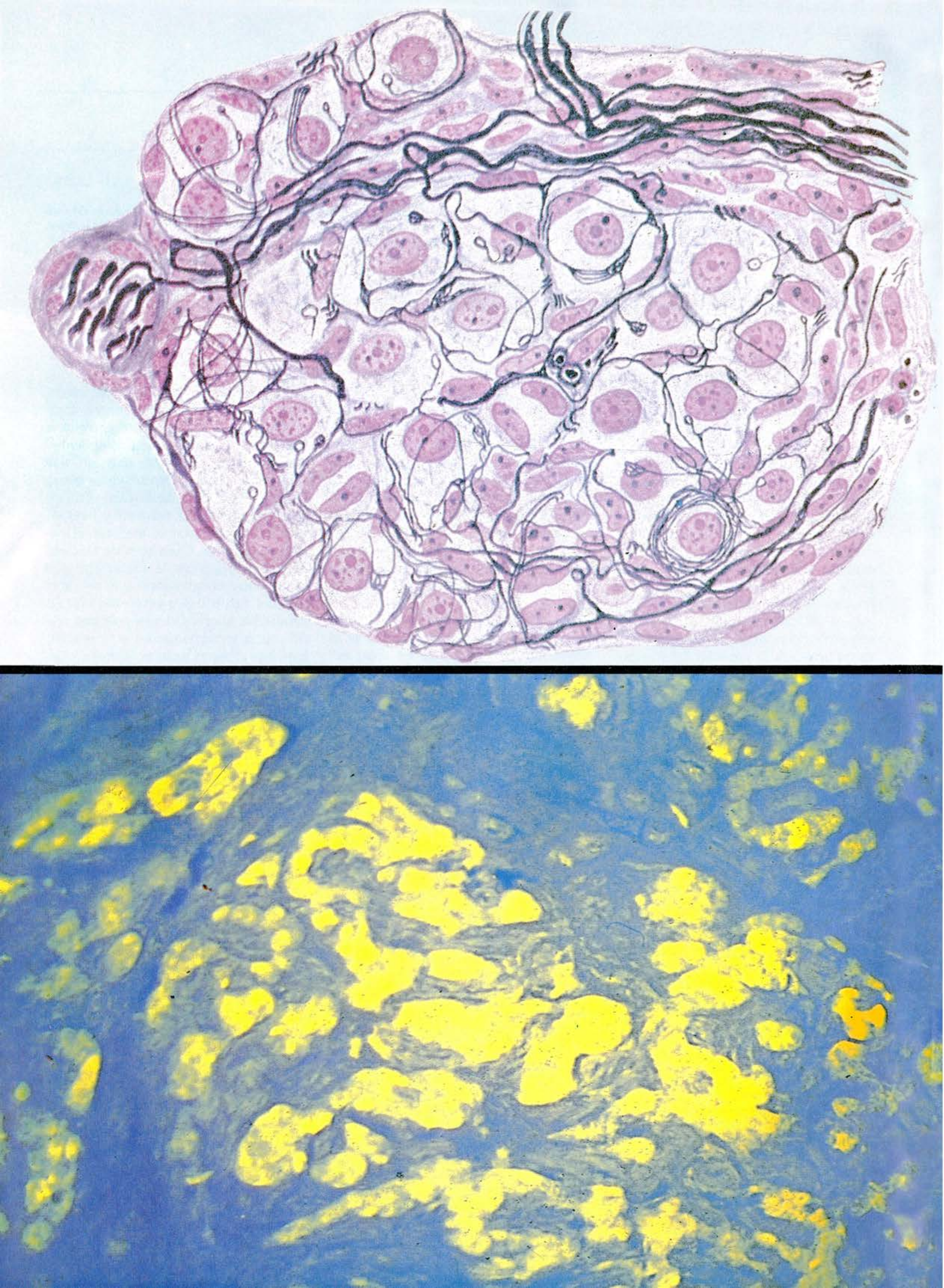
El sistema de Lilio confina la fecha del equinoccio vernal dentro de límites bastante estrechos. La fecha puede ser el 21 de marzo, el 20 de marzo o incluso el 19 de marzo, aunque no ha caído en el 19 de este mes desde, aproximadamente, los comienzos de este siglo. Pese a lo que se dice en muchos manuales de astronomía, el equinoccio de primavera, como resultado de las reglas gregorianas de intercalación, coincide más frecuentemente con el 20 que con el 21 de marzo.

Una consecuencia de la intercalación gregoriana es que todas las fechas en el calendario se repiten en un ciclo de, precisamente, 146.097 días, que equivalen a 400 años gregorianos. Dicho de otro modo, cada fecha de 1583 se repetirá en 1983; lo mismo se aplica a 1584 y 1984, y así sucesivamente en múltiplos de 400 años. El calendario completa su primer gran ciclo el 15 de octubre de este año.

Lilio y Clavius tuvieron éxito donde otros habían fracasado. El calendario gregoriano proporciona un compromiso muy satisfactorio entre una exactitud esencial y una simplicidad muy deseada. Durante más de 800 años, hombres tan competentes como Roger Bacon, Nicolás de Cusa, Regiomontano, Johannes Schöner y Paul de Middleburg intentaron mejorar el calendario juliano. Cada hombre de ciencia había registrado la disparidad creciente entre el calendario y el sol, mas por una razón o por otra –incluyendo entre ellas las rivalidades políticas, indiferencia gubernamental o la muerte prematura– no se puso ningún remedio a esta situación hasta que, en 1572, un antiguo profesor de derecho de Bolonia, llamado Ugo Buoncompagni, se convirtió en el Papa Gregorio XIII.



LOS CALENDARIOS JULIANO Y GREGORIANO aparecen, aquí, comparados en relación al desplazamiento de la fecha del equinoccio vernal. En el juliano este desplazamiento (*en color*) era bastante rápido. En el calendario gregoriano es más lento (*en negro*), porque el calendario se aproxima mejor al año trópico. El año trópico era equivalente al año gregoriano de 365,2425 hacia el año 3000 a. de C. Desde entonces dicha divergencia se ha acelerado.



El cuerpo carotídeo

Un pequeño ovillo de vasos sanguíneos, células glómicas y terminaciones nerviosas detecta cambios de oxigenación, anhídrido carbónico, acidez, temperatura, flujo y osmolaridad sanguínea. Su ubicación es variable

Carlos Eyzaguirre y Patricio Zapata

El cuerpo carotídeo es una pequeña estructura, situada en la bifurcación de la arteria carótida. Pesa aproximadamente 12 miligramos en humanos adultos y sólo 1 miligramo en el gato, animal de experimentación empleado frecuentemente en su estudio. El pequeño tamaño de este órgano se refleja en los nombres con que se lo ha bautizado a lo largo de su historia: *minutium, exiguum, parvum*, etcétera. Se encuentra vecino al "seno carotídeo", dilatación restringida de la pared arterial de la bifurcación carotídea. La ubicación exacta del cuerpo carotídeo varía de una especie a otra. En la conformación más corriente, la del gato adulto, la carótida común se continúa con la carótida externa y la carótida interna (a diferencia del hombre) es sólo un remanente embrionario; en el ángulo de esta bifurcación hallamos una arteria que se bifurca de inmediato en la faríngea ascendente y la occipital; es en el seno de esta segunda bifurcación donde se ubica el cuerpo carotídeo, con forma de pera cuyo tallo corresponde al nervio carotídeo (nervio sinusal de Hering), una de las ramas del IX par craneano, el nervio glosofaríngeo. Parte de las fibras del nervio carotídeo (sinusal) pasan de largo por el cuerpo carotídeo para ir a terminar al seno carotídeo. Ambas estructuras reciben además inervación simpática desde el ganglio cervical superior a través de los nervios ganglio-glomerulares (nervio carotídeo interno).

Dos hechos colocan de inmediato al cuerpo carotídeo en lugar destacado: es el órgano que, en relación a su masa, recibe el mayor aporte sanguíneo del organismo (2 litros por minuto y por 100 gramos de tejido) y que tiene un consumo metabólico extraordinariamente alto (10 mililitros de oxígeno por minuto y por 100 gramos de tejido). Ello indica que estamos en presencia de una maquinaria metabólica acelerada y exigente.

La ilustración superior de la página siguiente refleja diferentes supuestos sobre la naturaleza del cuerpo carotídeo: ¿se trata de un ganglio, de una glándula o de un receptor sensorial? Corresponde a un español, Fernando de Castro, el plantear en 1928 la naturaleza sensorial del órgano, luego de un detenido estudio de su estructura. Los hallazgos de de Castro, publicados en francés en *Trabajos del Laboratorio de Investigaciones Biológicas de la Universidad de Madrid*, han influido poderosamente en las investigaciones posteriores acerca de la función del cuerpo carotídeo.

La observación al microscopio óptico de cortes histológicos del cuerpo carotídeo revela un órgano profusamente vascularizado (de allí el nombre de glomus carotídeo), donde el parénquima o elemento principal se ve reducido a islotes o glomérulos de células epitelioides (llamadas también principales, receptoras, específicas, glómicas o tipo I), envueltas por células capsulares (sustentaculares, envolventes, intersticiales, satélites o tipo II). Las tinciones argénticas, por otra parte, revelan gran profusión de fibras nerviosas entre los islotes y dentro de ellos. De Castro pudo observar las finas terminaciones nerviosas usando como colorante azul de metileno. Las células glómicas poseen intensa fluorescencia natural debido a su alto contenido de catecolaminas.

¿Qué datos aportó la microscopía electrónica? Las células glómicas tienen gran abundancia de granos electrodensos con halo claro, de un diámetro entre 70 y 180 nanómetros, similares a los observados en las células cromafines de la médula adrenal. Muestran, además, extensas zonas de aposición con terminaciones nerviosas, ricas en vesículas claras (diámetro de 40 a 70 nanómetros), muy similares a las vesículas sinápticas de las terminaciones nerviosas en sinapsis o uniones neuroefectoras. Las similitudes señaladas sugirieron que las terminaciones nerviosas serían eferentes o motoras para las células glómicas, las que a su vez tendrían una función endocrina, entregando sus productos (posiblemente catecolaminas) a la sangre. Este concepto pareció llegar a substanciarse cuando en 1967 se publicaron observaciones de que dichas terminaciones degeneraban si se seccionaba las raíces del nervio glosofaríngeo. Observaciones que contradecían lo que de Castro había publicado en 1928, en el sentido de que las fibras y terminaciones nerviosas del cuerpo carotídeo persistían luego de seccionar las raíces del nervio glosofaríngeo. Lo que indujo a de Castro a postular que el pericarion (soma) de estas neuronas estaría alojado distalmente a las raíces intracraneales, pero centralmente a su trayecto fuera del cráneo, ya que lesiones nerviosas extracraneanas provocaban la virtual denervación del cuerpo carotídeo.

La ubicación del pericarion debería entonces estar en los ganglios sensitivos (petroso de Andersch y superior de Ehrenritter) del glosofaríngeo, alojados en el agujero yugular, lo que implicaba que las terminaciones nerviosas

COMO UNA ESTRUCTURA PROFUSAMENTE vascularizada e inervada se nos presenta el cuerpo carotídeo. La observación al microscopio óptico de cortes histológicos nos revela un órgano vascularizado, donde el parénquima se ve reducido a islotes o glomérulos de células epitelioides (llamadas también principales, receptoras, específicas, glómicas), envueltas por células capsulares. Las tinciones argénticas, por otra parte, revelan gran profusión de fibras nerviosas entre islotes y en el interior de éstos: arriba, cuerpo carotídeo fijado por perfusión, donde se destacan las células glómicas y su inervación (cortesía de Fernando de Castro). La microscopía de fluorescencia (*abajo*) pone de relieve la intensa fluorescencia natural de las células glómicas debido a su alto contenido en catecolaminas (cortesía de G. Muratori).

1743	TAUBE	Ganglion minutium
1762	HALLER	Ganglion exiguum
1772	NEUBAUER	Ganglion parvum
1775	ANDERSCH	Gangliolum intercaroticum
1833	MAYER	Ganglion intercaroticum
1862	LUSCHKA	Glandula carotica
1865	ARNOLD	Glomeruli arteriosi intercarotici
1891	MARCHAND	Nodulus caroticus
1892	RIEFFEL	Corpuscule retro-carotidien
1898	STÖHR	Glomus caroticum
1900	KOHN	Paraganglion intercaroticum
1904	FUNKE	Carotid body

HISTORIA DEL CUERPO CAROTIDEO a través de los distintos nombres con que se le conoció, desde mediados del siglo XVIII en que Taube se fija en la pequeñez de su estructura –como sus inmediatos seguidores– hasta el asentamiento de la expresión "cuerpo carotídeo" que le dio Funke, ya en el XX.

del cuerpo carotídeo eran sensitivas y no motoras. Varios autores, usando la electromicroscopía y la electrofisiología (Hess y Zapata, 1972; Nishi y Stensaas, 1974; McDonald y Mitchel, 1975), han confirmado posteriormente las observaciones de de Castro, demostrando que la inervación de las células glómicas es sensitiva, ya que persisten tras la sección intracraneana de las raíces del glosofaríngeo. Por otra parte, la inyección de aminoácidos radiactivos en el ganglio petroso permitió comprobar el transporte de estos trazadores hasta las terminaciones nerviosas que hacían contacto con las células glómicas.

De lo anterior se concluye que las terminaciones nerviosas sobre las células glómicas son sensoriales. A veces tienen apariencia de un cáliz que tiende a envolver la célula glómica, otras de un botón terminal que tiende a invaginarse en ella; ambos tipos de terminaciones y varias de forma intermedia proceden de la misma fibra sensitiva. La zona de aposición epitelio-neural puede ser bastante estrecha (de 20 a 30 nanómetros) y extensa, pero no se ha descrito fusión de membranas ('gap junctions'). La abundancia de vesículas sinaptoideas y mitocondrias en las terminaciones sensoriales del nervio carotídeo se observa también en mayor o menor grado en otros receptores, como corpúsculos gustativos y de Pacini, terminaciones anulo-espirales de los husos musculares, etcétera. Debe mencionarse, además, que se ha descrito la presencia de estas vesículas claras y sinaptoideas en el citoplasma de las células glómicas.

La aposición epitelio-neural podría constituir una sinapsis química recepto-sensorial que posee vesículas sinaptoideas a ambos lados de la unión (aunque más abundantes en la terminación nerviosa) y granos densos de halo claro muy abundantes en el lado epitelioide (excepcionales en las terminaciones nerviosas). Esto último ha venido a constituirse en criterio para reconocer prolongaciones citoplasmáticas de las células glómicas, que de otro modo podrían confundirse con terminaciones nerviosas. Además, se han encontrado uniones o sinapsis entre células glómicas vecinas. En cuanto a las sinapsis entre células glómicas y neuronas sensitivas, se ha planteado la posibilidad de que algunas de ellas sean sinapsis recíprocas. Estas últimas podrían participar tanto en el proceso de generación de impulsos sensoriales, como en procesos de retroacción sobre la célula glómica.

Se han realizado esfuerzos importantes para reconstruir tridimensionalmente las células glómicas y las terminaciones nerviosas del cuerpo carotídeo. Estas observaciones sugieren que el tamaño de la unidad quimiosensorial, definida como el número de células glómicas inervado por una misma fibra sensorial, varía de 10 a 20. El hecho de que la neurona sensorial no sólo presente ramificaciones terminales dentro del cuerpo carotídeo, sino también durante el trayecto extraglomerular del nervio carotídeo, sugiere la confluencia de impulsos sensoriales desde distintas fibras sobre un mismo axón. Esto podría hacer que la frecuencia de descarga de unidades sensoriales a nivel del proceso central de la neurona (raíces intracraneanas) fuera mayor que la ob-

a. Presencia del transmisor sospechoso.

b. Presencia de enzimas sintetizantes.

c. Presencia de sistemas de inactivación.

d. Presencia o incorporación de precursores metabólicos.

e. Menor eficiencia sináptica por inhibición de sistemas de síntesis, incorporación o almacenamiento.

f. Evocación de efectos post-sinápticos similares.

g. Potenciación de efectos post-sinápticos por inhibición de sistemas inactivadores.

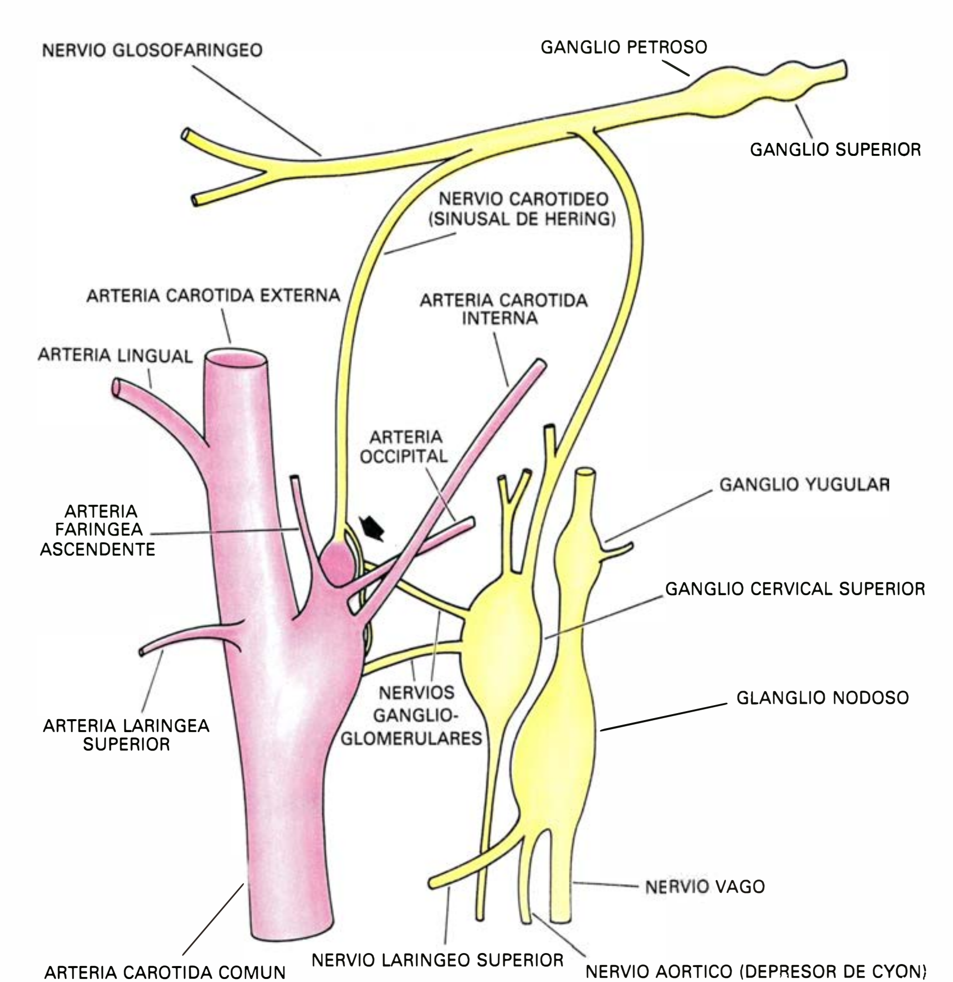
h. Bloqueo farmacológico de efectos post-sinápticos.

i. Liberación del transmisor.

DISTINTOS CRITERIOS SEGUIDOS para el reconocimiento de transmisores sinápticos. En la confección de ese listado se recogen proposiciones aportadas por diferentes autores. Quizá llame la atención del lector la omisión de un criterio: el que la transmisión sináptica puede interrumpirse en ausencia de iones calcio externos o de exceso de iones magnesio. Se trata, empero, de un criterio válido para cualquier proceso de transmisión química a través de una sinapsis, mas no ayuda a identicar el mediador.

tenida en registros unitarios a nivel del nervio carotídeo. Sin embargo, el impulso que se propaga desde una rama no sólo viajará ortodrómicamente hacia el pericarion y proceso central, sino también antidrómicamente hacia otras ramas, pudiendo entrar en colisión con los impulsos que se originan en sus respectivas terminaciones. Tales colisiones harán que el número de impulsos que se propaga hacia el sistema nervioso central sea menor que la simple suma de los originados en la periferia. Finalmente, la invasión antidrómica de las terminaciones sensitivas podría desencadenar en ella la liberación de sustancias químicas o transmisores que actúen sobre las células glómicas.

De lo anterior se desprende que si bien las células glómicas están inervadas por neuronas sensitivas, la organización de las unidades quimiosensoriales y la presencia de sinapsis recíprocas pueden permitir la transmisión de señales desde la fibra nerviosa hacia la célula glómica.



UBICACION E INERVACION del cuerpo carotídeo, estructura minúscula situada en la bifurcación de la arteria carótida (*flecha*). Se encuentra en la vecindad del seno carotídeo, dilatación restringida de la pared arterial de la bifurcación carotídea. La ubicación exacta del cuerpo carotídeo varía de una especie a otra. En la conformación más corriente, la del gato adulto, la carótida común se continúa con la externa.

La existencia de sinapsis recepto-sensoriales en el cuerpo carotídeo nos lleva a clasificarlo como receptor compuesto (o secundario); vale decir: un órgano en el que el proceso de detección de estímulos requiere la presencia de elementos neurales (terminaciones axónicas sensoriales) y células epitelioides preneurales. Otros receptores compuestos son los corpúsculos gustativos de la lengua y el órgano de Corti del oído. Todos ellos se contraponen a los receptores simples (o primarios), en los que la transducción del estímulo y generación de impulsos nerviosos se efectúa sin la mediación de células preneurales, careciendo, por supuesto, de sinapsis recepto-sensoriales. Es el caso del corpúsculo de Pacini, de los husos neuromusculares, de los órganos tendinosos de Golgi, de los barorreceptores del seno carotídeo y de los olfactorreceptores. En ambos tipos de receptores hay células accesorias no-neurales, pero ellas no parecen participar directamente en los procesos de transducción sensorial. Es posible que cumplan funciones de protección mecánica y metabólica de los elementos propiamente sensoriales.

El cuerpo carotídeo se ha clasificado como un receptor compuesto debido a la presencia de células glómicas formando sinapsis con las terminaciones nerviosas. También resulta difícil concebir que estas terminaciones sean capaces de detectar por sí solas estímulos tan diferentes como cambios de la concentración de hidrogeniones y de las presiones parciales de oxígeno y anhídrido carbónico. Si las células glómicas no participasen directamente en el proceso de transducción quimiosensorial, las fibras nerviosas que terminan en el cuerpo carotídeo tendrían que ser específicamente quimiosensibles. Sin embargo, se ha logrado reinervar funcionalmente el cuerpo carotídeo (o, lo que es lo mismo, recuperar la quimiorrecepción) con fibras nerviosas mecano-sensoriales del nervio laríngeo superior. Estas fibras no responden normalmente a los estímulos químicos. El éxito de esta reinervación heterotípica admite tres explicaciones: (1) que las fibras mecano-sensibles sufran una reespecificación tornándose quimiosensibles al llegar a contacto con las células glómicas; (2) que la transducción de los estímulos químicos siga a cargo de las células glómicas (las que por ella merecerían el nombre de receptoras), las que al liberar transmisores sinápticos logran excitar las fibras regeneradas en su vecindad; (3) una combinación de ambos factores.

La participación de las células glómicas en el proceso de quimiorrecepción ha sido puesta en duda por algunos autores. Mitchell, Sinha y McDonald, de la Universidad de California en San Francisco, han descrito que los neuromas formados por la regeneración del nervio carotídeo serían quimiosensibles; sin embargo, estos autores no extirpan los cuerpos carotídeos, por lo que no descartaron una posible reconexión de fibras regeneradas con estos órganos. Bingman, Kienecker y Knoche, de la Universidad de Münster, han descrito algo similar. Sin embargo, Leitner, Verna y Roumy observaron que, después de la criocoagulación del cuerpo carotídeo, el nervio carotídeo regenerado recuperaba su actividad quimiosensorial sólo si restablecía conexión con islotes de células glómicas que sobrevivían al procedimiento. Por nuestra parte (Zapata, Stensaas y Eyzaguirre), observamos que provocando degeneración del nervio carotídeo por atriciones a distancias variables del órgano, la reaparición de actividad quimiosensorial coincidía con el restablecimiento de contactos entre las fibras en

regeneración y las células glómicas. Finalmente, Smith y Mills, de la Universidad de Duke, observaron que los nervios carotídeos regenerados en bifurcaciones carotídeas a las que se ha extirpado el cuerpo carotídeo no readquieren quimiosensibilidad.

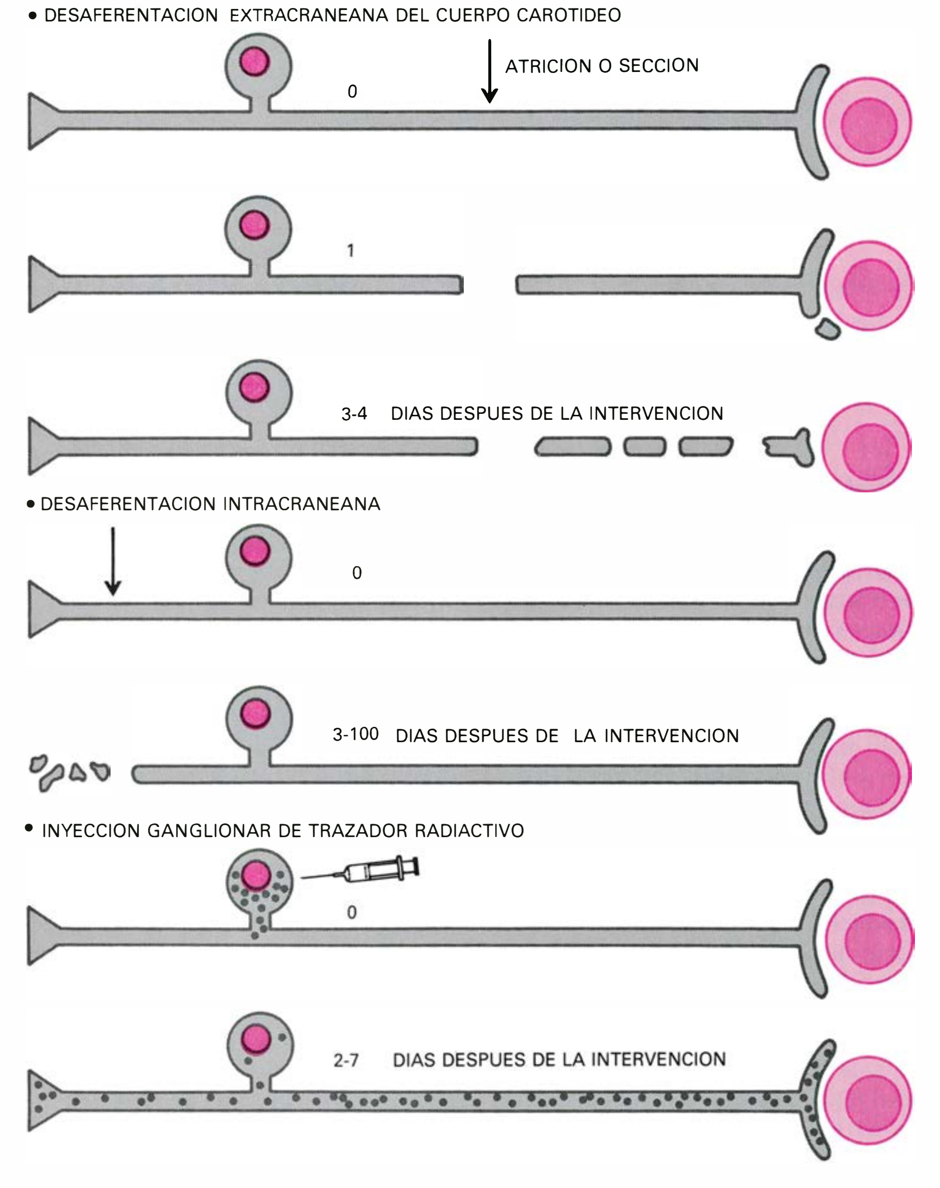
Estas tres series de observaciones indican que la conexión entre terminaciones nerviosas y células del cuerpo carotídeo es indispensable para restablecer la quimiorrecepción, aunque no permiten establecer cuál tipo de célula (glómica o capsular) es la comprometida en el proceso. Por otra parte, nuestros experimentos nos revelaron que la aparente normalización en las respuestas quimiosensoriales se obtenía antes del restablecimiento de las extensas sinapsis caliciarias, cuando los conos de crecimiento nervioso cercanos a las células glómicas eran aún escasos y estaban aún muy inmaduros. Ello nos muestra que la extensa unidad quimiosensorial normal, descrita anteriormente, contiene un alto factor de seguridad.

¿Cuáles son los estímulos para este órgano receptor de estructura compleja? La función del cuerpo carotídeo en la detección de cambios químicos de la sangre (Pa_{O_2}, Pa_{CO_2} y pHa) es bien conocida desde los clásicos estudios de C. Heymans, que le valieron el premio Nobel de fisiología y medicina en 1938. De hecho, el cuerpo carotídeo es un quimiorreceptor muy sensible no sólo a la hipoxia, hipercapnia y acidosis, sino también a una variada gama de agentes farmacológicos; de ellos, el cianuro (que provoca hipoxia citotóxica) y la nicotina (agente colinomimético) se han usado extensamente para explorar la función del cuerpo carotídeo. Respecto de los estímulos químicos naturales (Pa_{O_2}, Pa_{CO_2} y pHa), vale la pena recordar que la activación quimiosensorial evocada por aumentos progresivos del CO_2 en sangre obedece a una relación lineal. Pero la activación quimiosensorial, en respuesta a la disminución progresiva de la presión parcial del O_2 en sangre, presenta una relación hiperbólica o exponencial. La presentación simultánea de ambos tipos de estímulos revela una importante potenciación, vale decir, respuestas que superan ampliamente la simple interacción aditiva. Además, el CO_2 parece excitar directamente el cuerpo carotídeo y no sólo a través de cambios del pH extracelular, ya que aumentos del Pa_{CO_2} provocan estimulación quimiosensorial aun cuando se mantenga fijo el pHa.

La estructura de cuerpos carotídeos y aórticos es esencialmente la misma y sólo existen diferencias cuantitativas en su reactividad frente a algunos estímulos. La hipoxia hipoxémica, producida por una reducción en la presión parcial de oxígeno en la sangre arterial (Pa_{O_2}) debida a una disminución en la concentración de O_2 del aire inspirado o a baja presión atmosférica a gran altura, provoca excitación de los cuerpos carotídeos y aórticos. También se excitan por la caída del contenido de oxígeno en sangre arterial (Ca_{O_2}), producida al agregar monóxido de carbono (CO) al aire inspirado (el CO desplaza al O_2 a nivel de la hemoglobina del eritrocito), aunque se mantenga normal la Pa_{O_2}.

En 1959 se comenzó a emplear en el Departamento de Fisiología de la Universidad de Utah una técnica de estudio diferente, la superfusión in vitro del cuerpo carotídeo, que permitió replantearse interrogantes no resueltos por los estudios con la preparación in situ con circulación sanguínea y analizar nuevos problemas. La técnica consistía en extirpar el cuerpo y nervio carotídeo en animales anestesiados, lo que debía hacerse rápidamente luego de interrumpida la circulación carotídea, para colocarlos en un pequeño canal por el que fluía en forma continua una solución salina a la que se burbujeaba oxígeno y que los bañaba (superfundía) desde fuera. La preparación sobrevive en estas condiciones por muchas horas, lo que permite realizar en ella prolongados estudios bioquímicos, electrofisiológicos y farmacológicos.

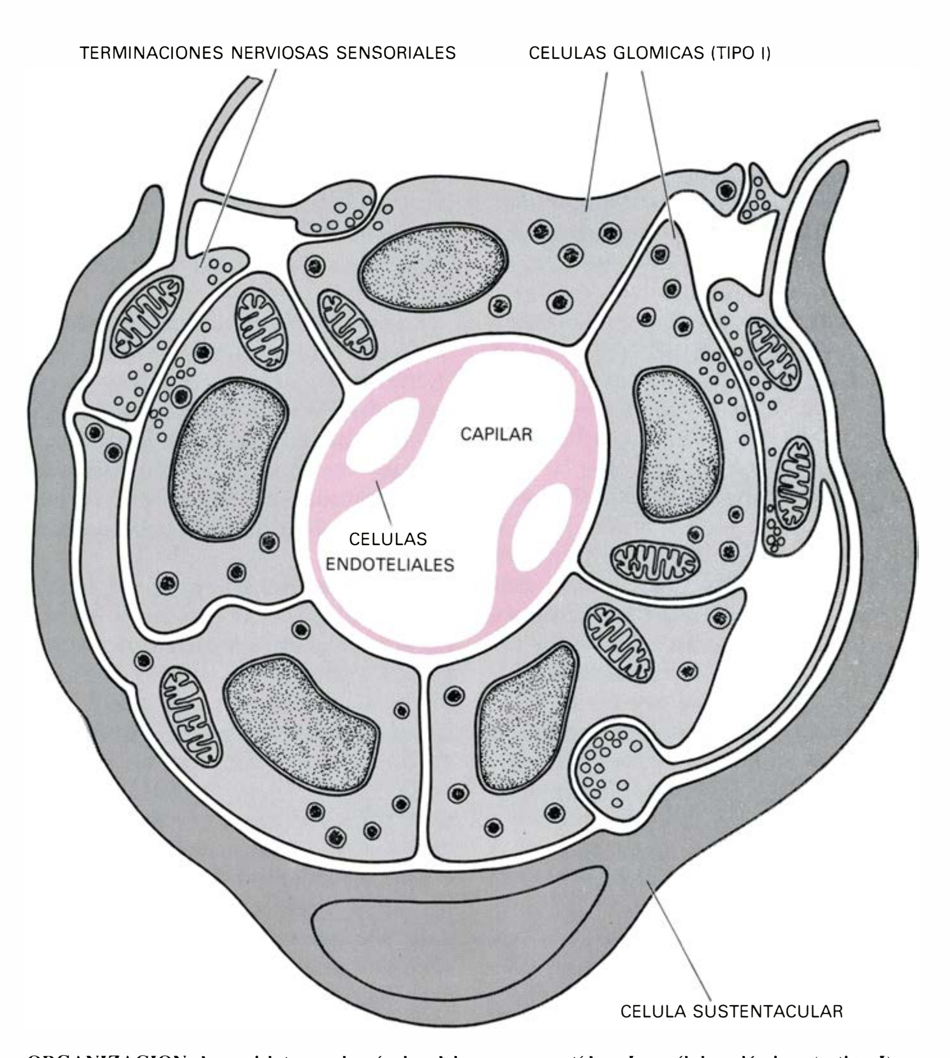
Empleando cuerpos carotídeos de gato superfundidos in vitro se ha confir-



LA INERVACION DE LAS CELULAS GLOMICAS es sensitiva, no motora, como ya había demostrado De Castro en 1928. Dicha inervación persiste tras la sección intracraneana de las raíces del glosofaríngeo. La inyección de aminoácidos radiactivos en el ganglio petroso permitió comprobar el transporte de los trazadores hasta las terminaciones nerviosas que ponían en contacto con las células glómicas.

mado la actividad quimiorreceptora en respuesta a hipoxia, hipercapnia y acidosis, como también la provocada por cianuro y nicotina. La misma preparación ha permitido estudiar modificaciones de la actividad quimiosensorial en respuesta a cambios de la temperatura del líquido de superfusión. Al elevar la temperatura, los cambios en la frecuencia de descarga de impulsos de una unidad quimiosensorial permiten distinguir un componente dinámico transitorio (sólo presente cuando el alza térmica es brusca) y un componente estático prolongado (determinado por el nivel térmico absoluto). Al registrar de todo el nervio, la frecuencia de impulsos quimiosensoriales a temperaturas bajas está constituida por impulsos infrecuentes en pocas fibras del nervio. Al subir la temperatura aumenta la frecuencia de descarga de cada fibra y se reclutan –entran en acción– fibras que hasta entonces estaban en silencio. La ilustración de la página 105 resume las frecuencias medias de impulsos nerviosos en relación a las temperaturas a que es sometido el cuerpo carotídeo. De este tipo de observaciones se desprende que el cuerpo carotídeo es un detector extremadamente sensible, preciso y de amplio rango para los cambios térmicos. Con estos datos se puede calcular el cociente térmico Q_{10} (razón entre una actividad fisicoquímica o biológica a cierta temperatura y a una temperatura inferior en 10 grados Celsius) o bien aplicar la ecuación de Arrhenius para calcular la energía de activación aparente (μ) de este órgano receptor. Los valores que se obtienen son excepcionalmente altos, lo que nos obliga a aceptar que estamos en presencia de un termorreceptor. McQueen y Eyzaguirre observaron que las fibras del nervio carotídeo in situ también respondían con gran precisión a los cambios de temperatura de la sangre circulante por la carótida. Tanto in situ como in vitro se observa que las fibras termosensibles son las mismas fibras quimiosensibles a cambios de gases y acidez. En cambio, las fibras mecanosensoriales, que inervan el seno carotídeo adyacente, modifican poco su frecuencia de descarga ante cambios de temperatura.

Experimentos realizados en la Universidad de Utah en colaboración con Roberto Gallego y Luis Monti-Bloch, obligaron a plantear el interrogante de si el cuerpo carotídeo no sería también un osmorreceptor. Trabajando con preparaciones superfundidas in vitro –con lo que se evita efectos vasculares colaterales– se observó que las soluciones hiperosmóticas (+ 15 a + 31 por ciento) producían un aumento de la frecuencia de descarga quimiosensorial generada en el cuerpo carotídeo, mientras que la superfusión con soluciones hipoosmóticas (− 8 a − 23 por ciento) producía el efecto opuesto. Es posible, por tanto, que el cuerpo carotídeo también participe en los procesos fisiológicos de detección de osmolaridad.



ORGANIZACION de un islote, o glomérulo, del cuerpo carotídeo. Las células glómicas (o tipo I) se hallan envueltas por células capsulares (llamadas sustentaculares, intersticiales, satélites o tipo II).

Años atrás, trabajando juntos con la preparación de cuerpo carotídeo superfundido in vitro, hubo un tipo de reacción que nos sorprendió. Estábamos habituados a observar que la detención súbita del flujo a través de la canaleta de superfusión producía un aumento sumamente rápido e intenso de la frecuencia de descarga de las fibras quimiosensoriales; esta frecuencia elevadísima se mantenía en el tiempo, a menos que se prolongase excesivamente la detención del flujo. Lo que nos llamó la atención fue que la frecuencia de descarga se reducía transitoriamente mientras se agitaba el líquido vecino al cuerpo carotídeo. En cambio, si se había establecido un régimen estacionario con una frecuencia quimiosensorial baja (lograda al superfundir el cuerpo carotídeo con una solución salina burbujeada con 100 por ciento de O_2) la agitación producía un aumento transitorio de la frecuencia de descargas. Nos pareció que el flujo de superfusión modificaba la actividad quimiosensorial.

¿Es, entonces, el cuerpo carotídeo también un fluxorreceptor? Este problema es difícil de resolver al trabajar con la preparación in situ, ya que la circulación a través del cuerpo carotídeo parece presentar autorregulación y, además, las fibras simpáticas procedentes del ganglio cervical superior controlan la musculatura vascular del órgano y particularmente el cierre o apertura de anastomosis arteriovenosas. Así, en colaboración con Belmonte, trabajando en la Universidad de Utah, se mostró que al estimular las fibras ganglioglomerulares las alzas en la frecuencia de descarga de las fibras quimiosensoriales del nervio carotídeo eran muy marcadas cuando el cuerpo carotídeo era perfundido (líquido que fluye a presión por el interior de los vasos ca-

rotídeos). Este aumento se reducía o desaparecía cuando el mismo órgano era superfundido (líquido que baña por fuera el cuerpo carotídeo).

Existe un problema respecto a la circulación sanguínea a través del parénquima de este órgano. Investigaciones realizadas en el Instituto Max-Planck de Dortmund (donde trabajan Lübbers, Acker y colaboradores) llevaron a plantear la posibilidad de que los islotes de células glómicas estuvieran alimentados por capilares a través de los cuales sólo fluye plasma sanguíneo, mientras los glóbulos se desplazarían por vasos extraparenquimatosos, tipo anastomosis arterio-venosas. Las observaciones referidas se realizaron en gatos. Sin embargo, los estudios electromicroscópicos de cuerpos carotídeos de conejos realizados por Verna (en Burdeos) plantearían la posibilidad de que el parénquima de este órgano recibiese sangre "completa". El problema tiene gran importancia, dada la capacidad de transporte de O_2 de la hemoglobina contenida en los eritrocitos.

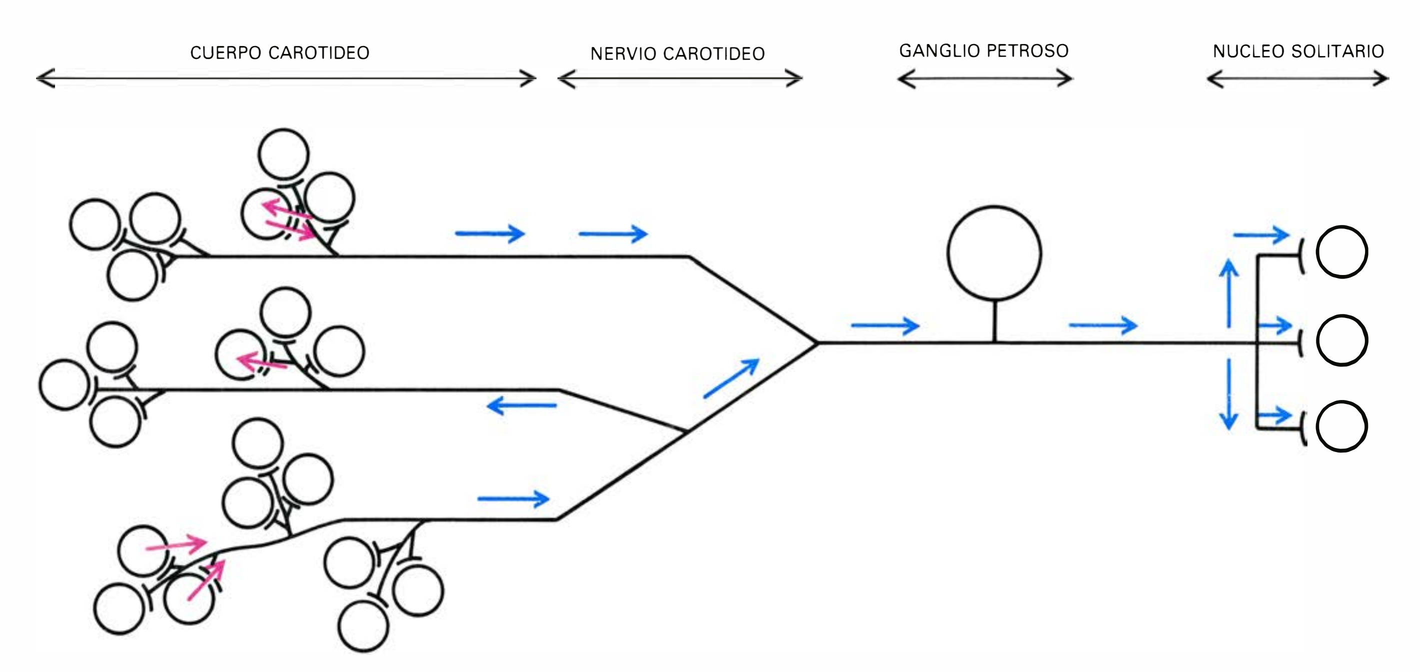
Cabe preguntarse, finalmente, si la presión arterial modifica la actividad de los quimiorreceptores del cuerpo carotídeo. Numerosos investigadores han mostrado que la actividad quimiosensorial no se modifica dentro de un amplio rango de regímenes de presión arterial; sólo cuando ésta cae aguda y profundamente, como en el caso de una hemorragia, se observa un incremento de la actividad quimiosensorial, en parte producido por cambios reflejos de la inervación simpática del cuerpo carotídeo. Pareciera que la presión arterial afectaría la actividad quimiosensorial sólo cuando llega a producir vasoconstricción local, cambios de flujo o hipoxia del cuerpo carotídeo. Esta insensibilidad a la presión arterial difiere radicalmente de la fidelidad con que las terminaciones barorreceptoras del seno carotídeo la detectan.

En conclusión, el cuerpo carotídeo se comporta como quimiorreceptor, termorreceptor, osmorreceptor y fluxoreceptor. Esta información la envía a través de las mismas fibras nerviosas, a las que habíamos llamado unidades quimiosensoriales. Son, pues, receptores y unidades sensitivas polimodales.

El cuerpo carotídeo es un órgano receptor sobre el cual se ha ensayado la más extensa variedad de fármacos. El análisis de los efectos de drogas se hizo inicialmente estudiando los cambios respiratorios evocados desde el cuerpo carotídeo; posteriormente, registrando del nervio carotídeo y estudiando las variaciones de la frecuencia de impulsos quimiosensoriales.

Un primer grupo de drogas es el de los venenos del metabolismo oxidativo, entre los que se destacan los inhibidores de la citocromo oxidasa –como es el caso del cianuro– y los inhibidores de la fosforilación oxidativa –como es el caso del 2,4-dinitrofenol–. Desde temprano se atribuyó sus efectos a una hipoxia citotóxica, lo que se ha confirmado más recientemente al observar que la estimulación quimiosensorial provocada por estas drogas aumenta durante hipoxia (pero no hipercapnia) y disminuye durante hiperoxia. Estando intactos los nervios carotídeos, la excitación quimiosensorial se manifiesta por un incremento reflejo de la ventilación pulmonar. La respuesta hiperventilatoria a la inyección intravenosa de cianuro es tan específica que se ha constituido en la prueba habitual para determinar la persistencia o interrupción de la vía neural entre cuerpos carotídeos y centros respiratorios bulbares. En casi todos los mamíferos es mayor la participación de los cuerpos carotídeos que la de los aórticos en el origen de la respuesta.

Un segundo grupo de drogas que provoca excitación quimiosensorial es el de los agentes colinomiméticos. Sin embargo, hay marcadas diferencias entre las distintas especies animales. En el gato, la inyección intracarotídea de acetilcolina provoca una descarga fugaz de las fibras quimiosensoriales; también in vitro, la adición de la droga al líquido de superfusión estimula el cuerpo carotídeo. En el conejo, tanto in situ como in vitro, la acetilcolina deprime la descarga quimiosensorial. El tratamiento de estas preparaciones con agentes anticolinoesterásicos, como la eserina o prostigmina, provoca intensa sensibilización y prolongación a los efectos de administraciones posteriores de acetilcolina. Esta observación indica la existencia de colinoesterasa en el cuerpo carotídeo, enzima que separa rápidamente el radical acetilo de la colina, siendo la colina mucho menos efectiva que la acetilcolina como estimulante del cuerpo carotídeo.

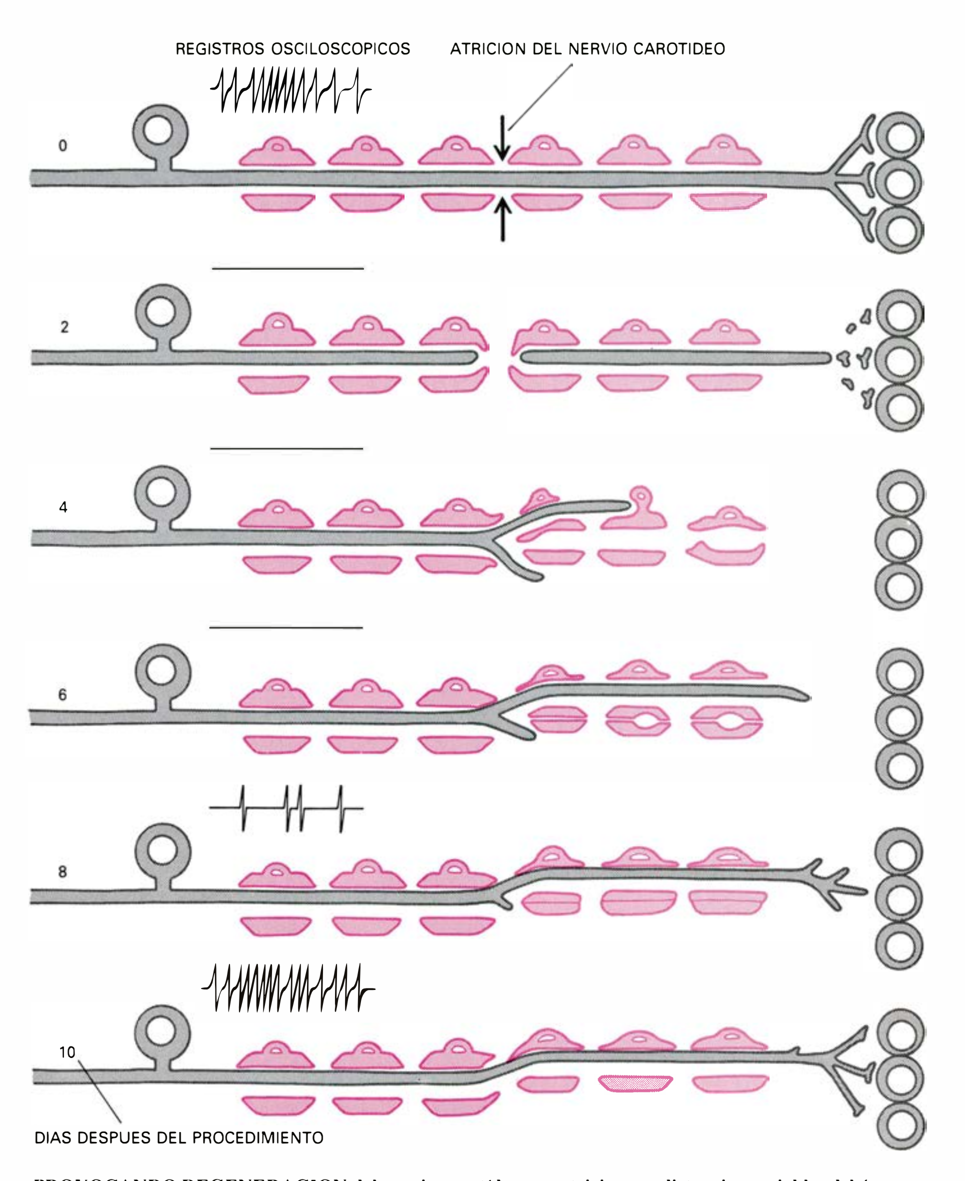


POR UNIDAD QUIMIOSENSORIAL se entiende el número de células glómicas inervado por una misma fibra sensorial, y su tamaño varía de 10 a 20 células. Las flechas rojas indican sentidos de transferencia de información sináptica recepto-neuronal; las azules, de propagación de impulsos nerviosos.

Existen numerosos estudios tendentes a caracterizar farmacológicamente los colinorreceptores del cuerpo carotídeo. Los agonistas nicotínicos, como la nicotina misma y la suberilcolina, producen excitación quimiosensorial en el gato y en el conejo, la que es reducida y eventualmente abolida luego de la administración de anticolinérgicos nicotínicos como la mecamilamina, hexametonio y *d*-tubocurarina. Los agentes muscarínicos, como la pilocarpina y el betanecol, prácticamente no producen efectos sobre los quimiorreceptores carotídeos del gato. Sin embargo, deprimen la actividad quimiosensorial en el conejo, efecto bloqueado por la atropina. Estos experimentos nos indican que los quimiorreceptores carotídeos del gato poseen una preponderancia de receptores nicotínicos (N), mientras que en el conejo predominan los colinorreceptores muscarínicos (M), ya que la acetilcolina tiene un efecto principalmente inhibidor.

Trabajando en el Laboratorio de Neurobiología de la Universidad Católica de Chile, en colaboración con A. Zuazo y F. Llados, observamos que la hiperventilación provocada en el gato por inyecciones de nicotina o por inhalación de humo de cigarrillo (siempre que contenga nicotina) desaparece luego de la sección de los nervios carotídeos. El discreto aumento de la ventilación observado en humanos al término del consumo de cada cigarrillo podría deberse a la estimulación de los cuerpos carotídeos por los niveles de nicotina alcanzados en sangre.

Los agentes adrenomiméticos constituyen un tercer grupo de drogas. Inyecciones intraarteriales o intravenosas de adrenalina o noradrenalina provocan aumentos de la frecuencia de impulsos quimiosensoriales originados en el cuerpo carotídeo, pero tales efectos se asocian en dosis y curso temporal a las acciones vasopresoras concomitantes. Atribuimos esta excitación no a un efecto directo sobre el aparato quimiorreceptor, sino a la hipoxia y cambios de flujo en el cuerpo carotídeo provocados por vasoconstricción local. En preparaciones superfundidas in vitro, en las que no se producen modificaciones vasculares, hemos observado que la adrenalina y la noradrenalina no producen excitación quimiorreceptora. En dicha preparación, tampoco observamos efectos de la isoprenalina, la que tiene acción hipotensora cuando se la inyecta al animal. Lo anterior sugiere que no hay α- ni β- adrenoceptores que participen directamente en el proceso de quimiorrecepción del cuerpo carotídeo.



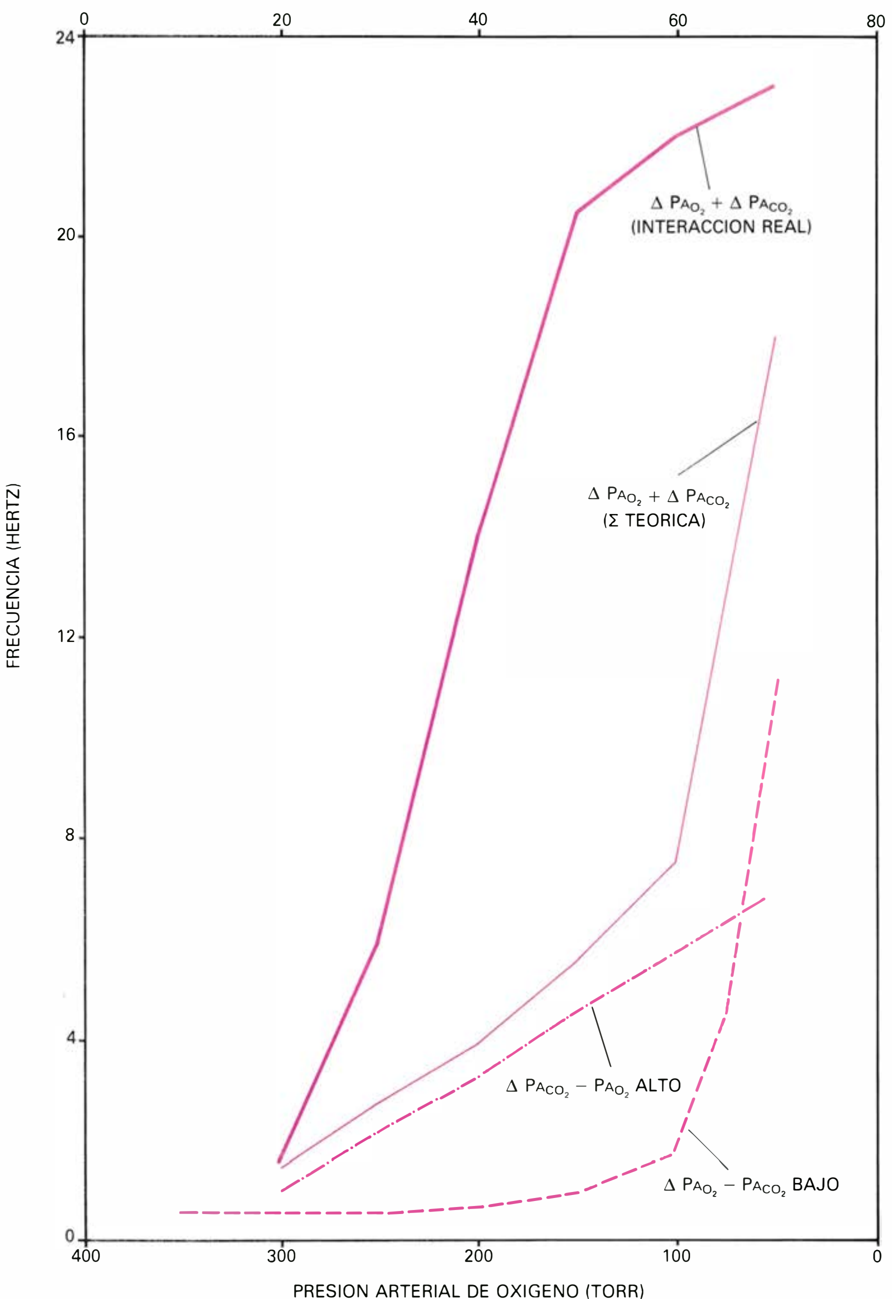
PROVOCANDO DEGENERACION del nervio carotídeo por atriciones a distancias variables del órgano, los autores, en unión con Stensaas, vieron que la reaparición de actividad quimiosensorial concidía con el restablecimiento de contactos entre las fibras en regeneración y las células glómicas. La aparente normalización en las respuestas quimiosensoriales se obtenía antes del restablecimiento de las sinapsis caliciarias, cuando los conos de crecimiento nervioso cercanos a las células glómicas eran aún escasos y estaban muy inmaduros todavía. Las flechas convergentes señalan atrición del nervio carotídeo; los números, días después del procedimiento. Los registros osciloscópicos son los obtenibles a esos intervalos. (La ilustración se funda en observaciones de los autores y Stensaas, en *Brain Research*, 113, p. 235-253, 1976.)

En el curso de los últimos años, y trabajando en distintas preparaciones (arteria renal, sistema nigroestriatal, etc.), se ha confirmado la existencia de dopaminoceptores, los que responden preferentemente a la dopamina, sustancia que en las células adrenérgicas y noradrenérgicas es el precursor inmediato de la noradrenalina. Hemos podido observar que la dopamina es capaz de producir inhibición quimiosensorial en el gato, tanto en preparaciones de cuerpos carotídeos in situ como superfundidos in vitro. En la primera, el efecto quimiosensorial se manifiesta con dosis de dopamina inferiores a aquéllas que producen efectos vasculares. Estudios realizados en el Laboratorio de Neurobiología de la Universidad Católica de Chile muestran que análogos de la dopamina, como la apomorfina y amantadina, también provocan inhibición quimiosensorial.

La inhibición provocada por la dopamina o sus análogos puede ser bloqueada si se administran antagonistas del tipo de las fenotiazinas (perfenazina y clorpromazina) o butirofenonas (haloperidol, espiroperidol). Los ergoloides y las ergolinas se comportan inicialmente como análogos y luego como antagonistas de la dopamina. Nos llamó la atención observar que posteriormente a la administración de perfenazina o de

espiroperidol, la administración subsecuente de dopamina, en dosis altas, provocaba ahora excitación quimiosensorial. Concluimos de ello la existencia de dopaminoceptores inhibidores (DA_i) y excitadores (DA_e) en el cuerpo carotídeo del gato. El efecto de la dopamina es diferente en otras especies. Por ejemplo, en el conejo la dopamina aplicada in vitro tiene especialmente un efecto excitador sobre la descarga quimiosensorial, mientras in situ tiene un efecto inhibidor. En el perro, la dopamina a dosis bajas inhibe la actividad quimiosensorial, pero puede aumentarla a dosis altas. La proporción de receptores DA_i y DA_e en distintos mamíferos podría explicar el que la dopamina produzca efectos predominantes inhibidores, excitadores o bifásicos.



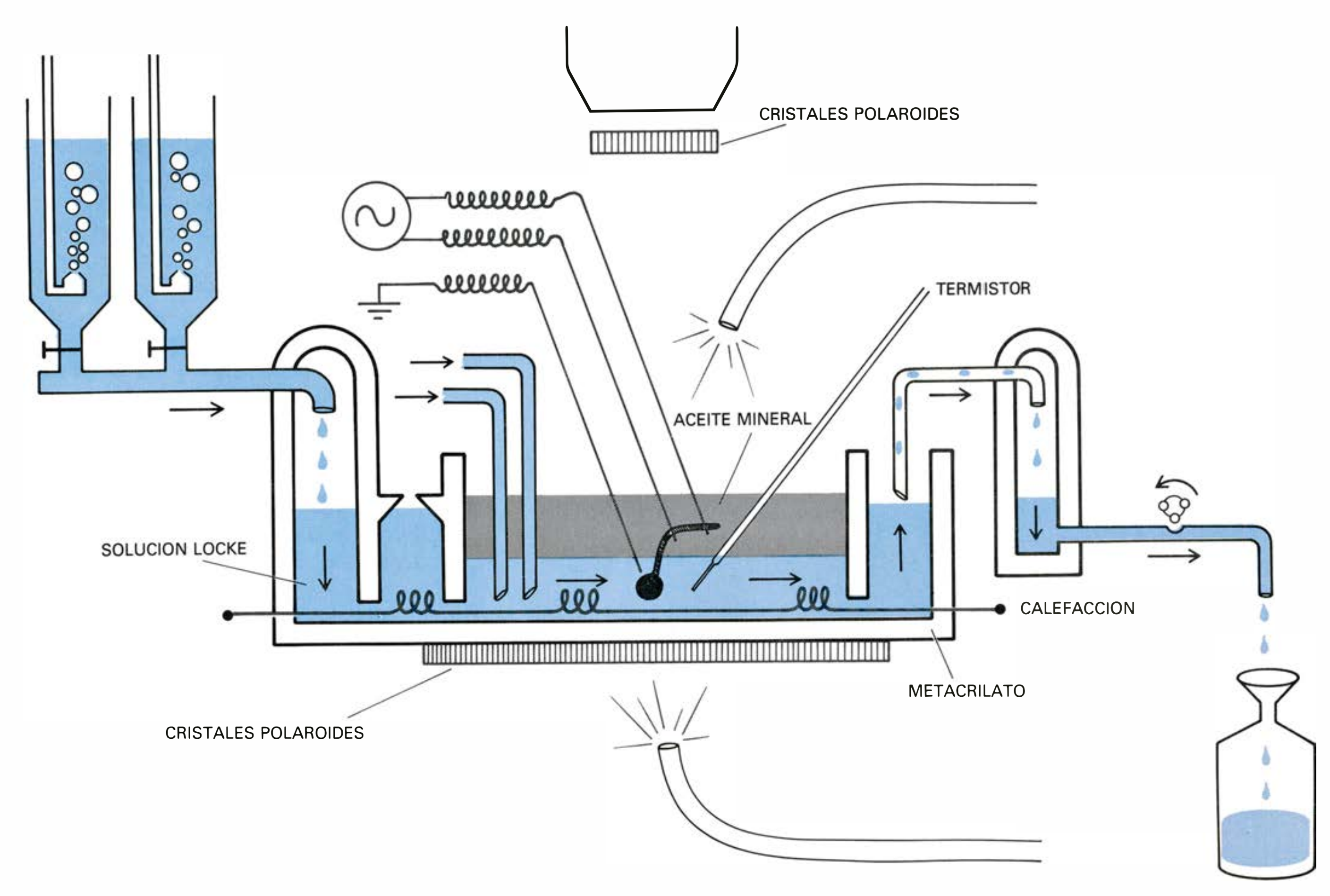


RELACION ENTRE FRECUENCIA DE DESCARGAS de una unidad quimiosensorial y niveles tensionales de gases en sangre arterial. La función del cuerpo carotídeo en la detección de cambios químicos en la sangre se conoce ya desde el primer tercio del siglo. Se trata, en efecto, de un órgano quimiorreceptor muy sensible, no sólo a la hipoxia, acidosis e hipercapnia, sino también a múltiples agentes farmacológicos. Respecto de los estímulos químicos naturales (presión arterial de oxígeno, de anhídrido carbónico, *p*H), vale la pena recordar que la activación quimiosensorial evocada por aumentos progresivos del CO_2 en sangre obedece a una relación lineal; pero la activación quimiosensorial en respuesta a la disminución progresiva de la presión parcial del O_2 en sangre presenta una relación hiperbólica o exponencial. La presentación simultánea de ambos tipos de estímulos revela una importante potenciación: respuestas que superan ampliamente la simple interacción aditiva. La interpretación del gráfico es la siguiente: línea roja clara, frente a cambios de oxigenación (indicados en abscisa inferior), manteniendo fijo Pa_{CO_2} bajo; línea roja interrumpida por puntos, cambios de CO_2 indicados en abscisa superior, manteniendo alta la tensión de O_2; línea clara, sumación teórica de los efectos anteriores; línea intensa, efectos reales de cambios simultáneos de gases arteriales. La presión se da en torr y, en hertz (ciclos por segundo), la frecuencia.

En el curso de los últimos años se han realizado en el Departamento de Fisiología de la Universidad de Utah intensos esfuerzos para analizar la participación de los diferentes componentes celulares del cuerpo carotídeo en el proceso de quimiorrecepción. Para ello ha sido necesario llegar a trabajar en cortes de tejido, intentando penetrar las células con microelectrodos bajo control microscópico con contraste por interferencia (óptica Nomarski), seguido en algunos casos de microinyección de colorantes, para su identificación.

El potencial de membrana (media geométrica de 19 milivolt) y la resistencia (media geométrica de 32 megaohm) de las células glómicas no son afectados significativamente por cambios en la concentración externa de ion potasio (K^+). La remoción del ion sodio (Na^+) externo produce hiperpolarización y aumento de la resistencia, mientras que la reducción del ion cloro (Cl^-) externo produce depolarización y caída de la resistencia. La hipotermia produce una reducción del potencial de membrana y de la resistencia de estas células, en tanto que la hipertermia tiene el efecto contrario. Esto parece indicar la presencia de bombas metabólicas que mantienen el equilibrio iónico entre las células glómicas y su ambiente. Conviene aquí destacar que es posible que los procesos de las células capsulares que envuelven a las células glómicas lleguen a establecer un medio extracelular inmediato de intercambio lento con el medio interno del organismo, aunque por el momento se trata sólo de una hipótesis.

Si se baña un cuerpo carotídeo con una solución que contiene 100 por ciento de N_2 o bien 6 por ciento de CO_2 a *p*H normal, algunas células se depolarizan, mientras otras se hiperpolarizan y un tercer grupo no cambia su potencial de membrana. Lo mismo ocurre cuando se aplica cianuro o acetilcolina. Sin embargo, la caída a 6,0 del *p*H del líquido de superfusión, agregando ácidos o bien burbujeándolo con CO_2 al 6 por ciento, produce una importante depolarización. La alcalinidad (*p*H 8,0) tiene un efecto hiperpolarizante. Finalmente, cabe destacar que la estimulación eléctrica del nervio carotídeo no produce cambios del potencial de membrana de la células glómicas, lo que se



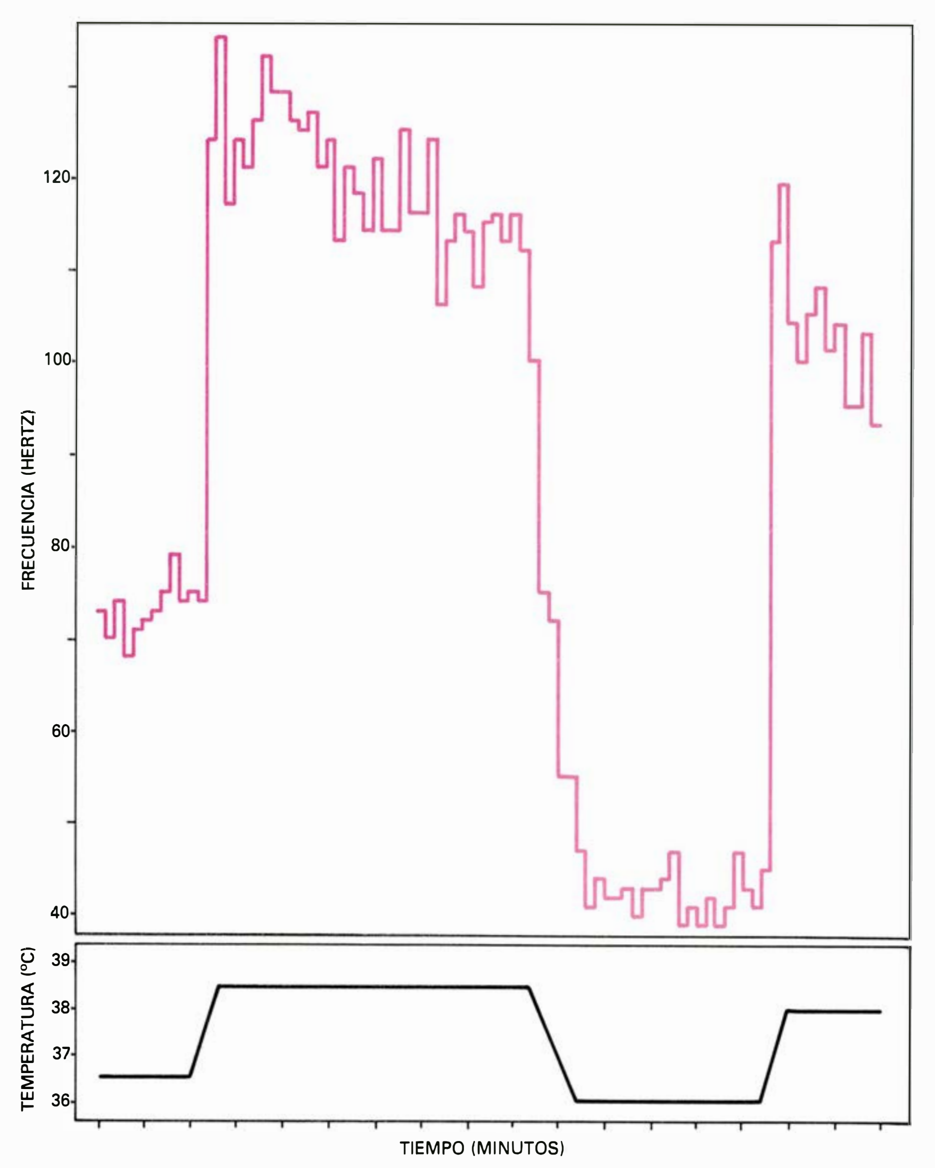
SISTEMA DE SUPERFUSION de cuerpo carotídeo in vitro y registro de actividad quimiosensorial. Esa técnica permitió plantearse de nuevo interrogantes no resueltos por los estudios realizados con preparaciones in situ con circulación sanguínea, amén de analizar nuevos problemas. La técnica consiste en extirpar el cuerpo y nervio carotídeo en animales anestesiados, lo que debe hacerse rápidamente luego de interrumpida la circulación carotídea, para colocarlos en un pequeño canal por donde fluye en forma continua una solución salina a la que se burbujea oxígeno y que los baña (superfunde) desde fuera. La iluminación es episcópica y diascópica con luz fría transmitida por tubos de fibra óptica; la observación, estereomicroscópica con luz polarizada. Las jeringas inyectan drogas en el canal. El circuito calefactor se controla por un tele-termistor. La preparación sobrevive horas en estas condiciones.

ha usado como argumento en contra de la existencia de fibras centrífugas en el nervio carotídeo que controlen la función de las células glómicas. Sin embargo, la inervación de las células glómicas parece condicionar algunas de sus respuestas, ya que corpúsculos denervados por más de dos semanas no presentan la típica respuesta depolarizante al enfriamiento.

Los impulsos nerviosos de las fibras quimiosensoriales deberían tener su origen en una depolarización de las terminaciones nerviosas (potencial generador). Por ello, se ha intentado, y logrado, registrar intracelularmente, pudiéndose penetrar aquellas terminaciones relativamente grandes (alrededor de 10 micrometros) demostradas por la microscopía electrónica. Los registros logrados de estas terminaciones han revelado potenciales depolarizantes espontáneos de 1 a 10 milivolt de amplitud y duración de 5 a 120 milisegundos, cuya frecuencia aumenta durante la estimulación. Algunos potenciales de acción se originan de la cima de los potenciales depolarizantes espontáneos, mientras otros no guardan relación alguna con ellos. Es posible que estos últimos correspondan a invasión antidrómica de la terminación por potenciales de acción generados en otra terminación de la misma unidad quimiosensorial.

Un medida más indirecta y global del proceso de generación de la actividad quimiosensorial puede obtenerse registrando in situ o in vitro con dos electrodos impolarizables, uno aplicado al polo neural del cuerpo carotídeo y otro a distancia en el nervio carotídeo, especialmente si entre ambos se interpone un espacio de sucrosa, aire o aceite mineral. Si en estas condiciones se aplica un estímulo natural o una droga excitadora se observa la aparición de un potencial depolarizante lento, local (no propagado), que no deja refractariedad (pueden sumarse los efectos de dos estímulos presentados en rápida sucesión) y resistente al bloqueo por tetrodotoxina y procaína; en suma, las características de los potenciales generadores descritos en otros órganos receptores. Se ha denominado potencial receptor masivo y descrito como depolarizante, ya que tiene igual sentido que la diferencia de potencial que aparece al aplicar K^+ al cuerpo carotídeo. Lo más probable es que el potencial receptor masivo corresponda a la suma algebraica de las depolarizaciones de un cierto número de terminaciones quimiosensoriales, generadoras de los impulsos respectivos. Sin embargo, es posible que también contribuyan a él cambios de potencial de las células glómicas, si hubiese acoplamiento eléctrico entre estas células y las terminaciones nerviosas.

Tenemos ahora la información necesaria para analizar el posible funcionamiento del cuerpo carotídeo como modelo de receptor complejo. Hipotéticamente, en un receptor de este tipo el proceso de transducción –en este caso la quimiorrecepción propiamente tal– correría a cargo de la célula glómica receptora y este proceso se evidenciaría por un cambio de su potencial de membrana. Los cambios eléctricos de la membrana de la célula receptora



REACTIVIDAD TERMICA de cuerpo carotídeo *in vitro*, superfundido con solución Locke modificada (*p*H 7,44 a 35 grados Celsius), equilibrada con O_2 al 100 por ciento, a flujo de 0,6 mililitros por minuto. El registro corresponde a todo el nervio carotídeo. Cada trazo horizontal respresenta la media de 12 segundos. Al elevar la temperatura, los cambios en la frecuencia de descarga de impulsos de una unidad quimiosensorial permiten distinguir un componente dinámico transitorio (sólo presente cuando el alza térmica es brusca) y un componente estático (que viene determinado por el nivel térmico absoluto).

podrían acompañarse de un incremento de la permeabilidad al Ca^{2+} y la entrada de este ion desencadenaría un proceso de exocitosis (acoplamiento excitación-secreción) de vesículas que contienen los paquetes de transmisores sinápticos. Serían estos transmisores (y no los estímulos) los que –por afinidad con receptores moleculares de la membrana de la terminación nerviosa– darían lugar a una depolarización local y graduada de la terminación sensorial (el llamado "potencial generador"), que llegada a cierto umbral generaría una ráfaga de impulsos (potenciales de acción) que se propagan a lo largo del axón.

Las pruebas experimentales obtenidas del registro extracelular de potenciales receptores masivos y del registro intracelular de terminaciones nerviosas coinciden globalmente con el modelo teórico. Por otra parte, las observaciones electromicroscópicas sobre las aposiciones epitelio-neurales y elementos vesiculares, junto a observaciones electrofisiológicas sobre los efectos de variacions del Ca^{2+} externo sobre la quimiorrecepción, son también compatibles con la presencia de una sinapsis sensorial en este órgano. Sin embargo, los registros intracelulares de células glómicas muestran respuestas (o falta de ellas) a los estímulos naturales muy difíciles de calzar con el modelo teórico. Si todos los estímulos fisiológicos hubiesen provocado depolarización de las células glómicas (o todos ellos hiperpolarización), el acuerdo habría sido perfecto, pero no es así. ¿Podrán entonces los estímulos naturales modificar la permeabilidad de algunas células glómicas sin cambios significativos del potencial de membrana y desencadenar la liberación de transmisores, mientras que otros lo hacen a través de cambios de potencial? Quizá la situación sea aún más heterogénea: ¿Podrá la detección de algunos estímulos ser mediada sinápticamente a través de liberación de transmisores de la célula glómica, mientras otros estímulos actúan directamente sobre una terminación nerviosa sensorial diferenciada por el contacto con la célula glómica?

Existe adecuada información que sugiere la intervención de un proceso de transmisión sináptica desde la célula glómica en la generación de actividad nerviosa quimiosensorial en respuesta a hipoxia y acidosis. Se ha demostrado que el cuerpo carotídeo contiene acetilcolina. Basados en analogía con sinapsis eferentes se pensó que debía estar alojada en las vesículas sinápticas tan prominentes en algunas terminaciones sensoriales; sin embargo, los nervios carotídeo y aórtico sólo tienen cantidades ínfimas de acetilcolina. La acetilcolina persiste en el cuerpo carotídeo crónicamente denervado; pruebas indirectas sugieren que estaría localizada en las células glómicas. Luego la transmisión sináptica sensorial podría ser colinérgica.

El reconocimiento de una sustancia como mediador sináptico no puede descansar en hallazgos aislados, sino que debe someterse a un detallado examen sobre un grupo de características que deben darse en conjunto. Varios autores han propuesto listas de estos criterios. La ilustración inferior de la página 96 resume una proposición actualizada. Puede llamar la atención la omisión de uno de ellos: el que la transmisión sináptica puede interrumpirse en ausencia de Ca^{2+} externo o exceso de Mg^{2+}; ello es un criterio válido para cualquier proceso de transmisión química a través de una sinapsis, pero no ayuda a identificar el mediador.

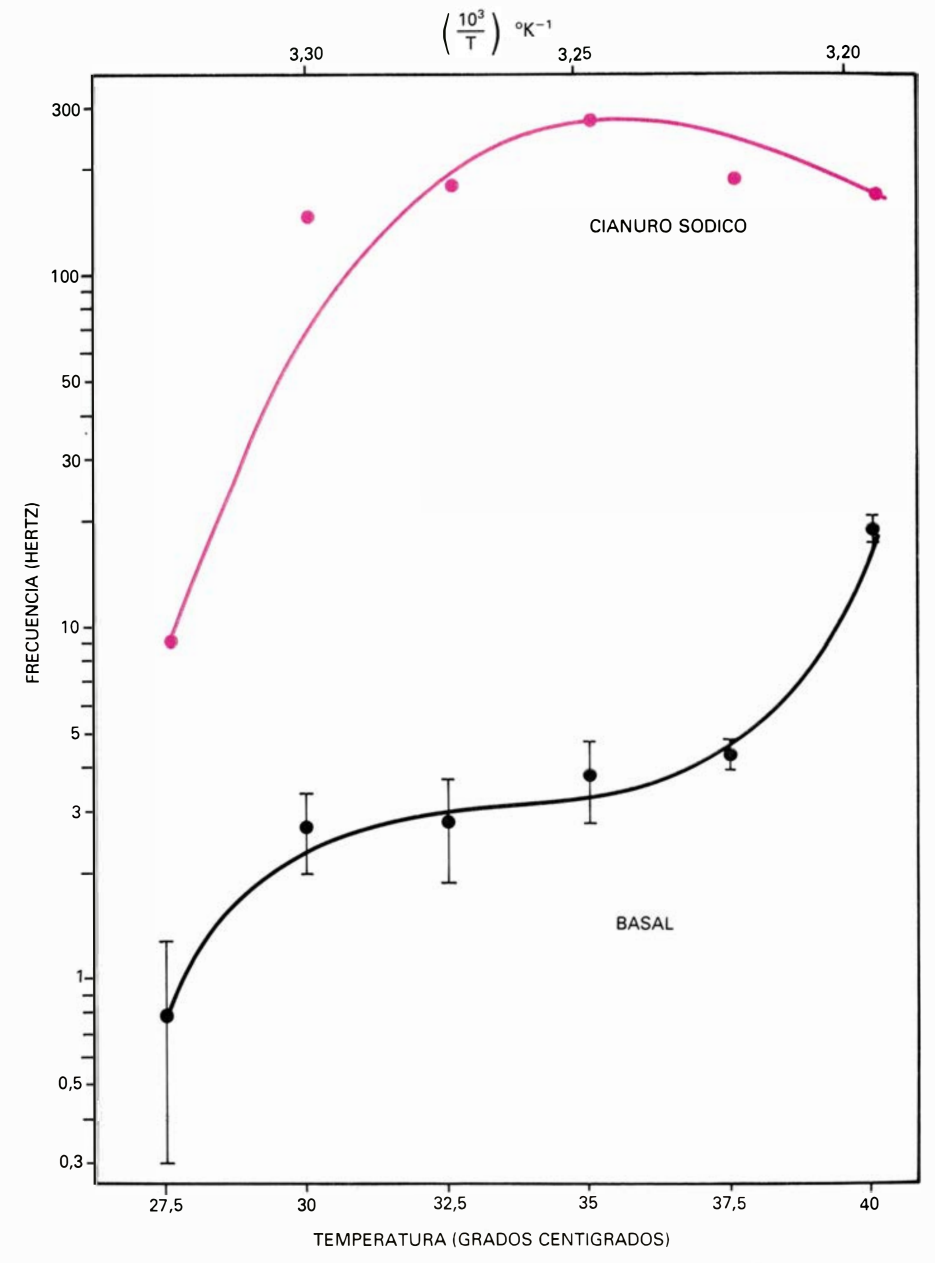
Aplicada la lista de criterios de la citada tabla a la candidatura de la acetilcolina como mediador excitador en la sinapsis sensorial del cuerpo carotídeo del gato, podemos observar que: (*a*) El parénquima del cuerpo carotídeo tiene un contenido relativamente alto de acetilcolina, si bien los valores obtenidos son todavía dispares. (*b*) El cuerpo carotídeo contiene el enzima sintetizante

colino-acetil-transferasa (colinoacetilasa). (*c*) Hay pruebas bioquímicas e histoquímicas de que el cuerpo carotídeo contiene los enzimas degradantes acetil-colinoesterasa (acetilcolino-hidrolasa) y butiril-colinoesterasa (pseudo-colinoesterasa). (*d*) Estudios autoradiográficos muestran alta afinidad en la incorporación del precursor, colina, a las células glómicas. (*e*) La reactividad de cuerpos carotídeos superfundidos in vitro se prolonga en presencia de colina agregada al líquido de superfusión. (*f*) La acetilcolina produce excitación quimiosensorial, al igual que la hipoxia, hipercapnia y acidosis. (*g*) El efecto excitador de la acetilcolina sobre el cuerpo carotídeo se intensifica y prolonga luego de inhibir la colinoesterasa con eserina; la eserinización del cuerpo carotídeo superfundido in vitro a veces aumenta la frecuencia basal de descarga quimiosensorial, potencia en algunos casos la respuesta al *p*H ácido, pero en otros la deprime (este último podría atribuirse a bloqueo por la propia acetilcolina que persiste demasiado sin ser inactivada).

Hay dos puntos más a observar. (*h*) Los efectos excitadores de la acetilcolina sobre el cuerpo carotídeo son bloqueados por mecamilamina, hexametonio, curare y atropina, aproximadamente en ese orden. La mayor objeción a la hipótesis colinérgica de transmisión sináptica en estos receptores se basa en que los estímulos naturales del cuerpo carotídeo siguen siendo efectivos después que agentes curarizantes bloquean los efectos de la acetilcolina. Existen observaciones discrepantes respecto al efecto de la mecamilamina sobre la respuesta a la hipoxia del cuerpo carotídeo in situ. En la preparación in vitro, hay disminución de las respuestas a acidosis e hipoxia durante el bloqueo con mecamilamina y hexametonio. (*i*) Metz, a la sazón en la Universidad de Carolina del Sur, pudo demostrar en 1969 la liberación de acetilcolina desde el cuerpo carotídeo del perro durante estimulación hipóxica y especialmente asfíxtica (hipoxia + hipercapnia).

Quisiéramos referir una observación más respecto al último punto. Si se hace pasar corriente eléctrica a través de un cuerpo carotídeo superfundido in vitro se registra de su nervio un aumento de la frecuencia de descarga de los impulsos quimiosensoriales que no se establece de inmediato y se prolonga mucho más allá del término del estímulo. Ahora bien, si el líquido que baña este cuerpo se aprovecha para superfundir un segundo cuerpo carotídeo [*véase la ilustración de la página 107*], se aprecia que tardíamente se produce un discreto aumento de la frecuencia de impulsos quimiosensoriales recogidos del nervio del segundo cuerpo carotídeo.

Esta técnica es una adaptación de la ingeniosa preparación de dos corazones de rana en serie utilizada por Otto Loewi para demostrar la liberación del "vagusstoff", que resultó ser la acetilcolina. En nuestra preparación, observamos que si se agregaba eserina al líquido de superfusión la respuesta del segundo cuerpo carotídeo se hacía mucho más notoria; respuesta que desaparecía al agregar mecamilamina al líquido de superfusión. Este experimento demuestra la liberación de una sustancia química desde el primer cuerpo carotídeo que ejerce un efecto excitador sobre el segundo, efecto potenciado por un anticolinoesterásico y bloqueado por un anticolinérgico. La sustancia colinérgica liberada procede del parénquima del cuerpo carotídeo, ya que el fenómeno también se presenta utilizan-



AL REGISTRAR TODO EL NERVIO, la frecuencia de impulsos quimiosensoriales a temperaturas bajas está constituida por impulsos infrecuentes en pocas fibras del nervio. Al subir la temperatura, aumenta la frecuencia de descarga de cada fibra y se reclutan –entran en acción– fibras que hasta entonces estaban en silencio. El gráfico recoge, resumidas, las frecuencias medias de impulsos nerviosos en relación con la temperatura a la que es sometido el cuerpo carotídeo. Se ha perfundido con solución Locke modificada (*p*H 7,44 a 35 grados Celsius) equilibrada con O_2 al cien por ciento, con un flujo de 1,5 mililitros por minuto (lo que explica las diferencias basales respecto a la ilustración precedente). En círculos negros, las medias aritméticas de actividad basal de 100 segundos; en rojos, el efecto máximo (media de 10 segundos) a inyecciones de cianuro sódico 20 microgramos en canal de superfusión. (Zapata: datos inéditos.)

do un primer cuerpo carotídeo crónicamente denervado.

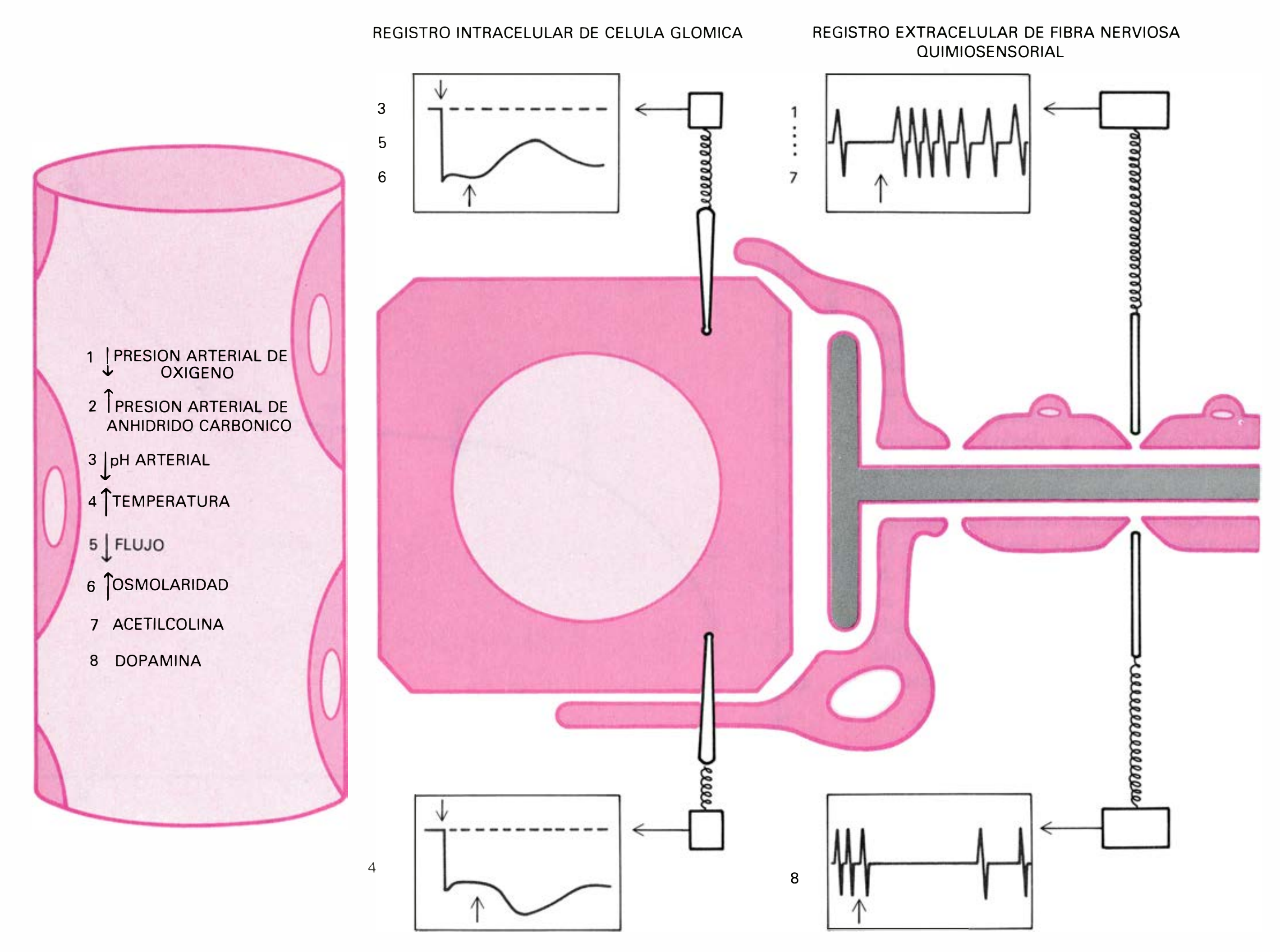
De la revisión anterior vemos que la acetilcolina cumple con casi todos los criterios para ser considerada mediador excitador en la sinapsis sensorial del cuerpo carotídeo. Sin embargo, la presencia de esta sinapsis colinérgica se hace insuficiente para explicar todos los procesos de quimiorrecepción. De hecho, en algunas especies animales, como el conejo, la acetilcolina tiene efectos fundamentalmente inhibidores.

La existencia de granos electrodensos con halo claro en las células epitelioides del cuerpo carotídeo hizo pensar que contuviesen catecolaminas, lo que vino a confirmarse posteriormente a través de varios métodos y numerosos experimentos. Conviene entonces aplicar la misma lista de criterios de la ilustración inferior de la página 96 frente a una posible sinapsis catecolaminérgica en cuerpo carotídeo. (*a*) El cuerpo carotídeo encierra un alto contenido de catecolaminas, que corresponde en su mayor proporción a dopamina: 63 por ciento en el gato, 77 por ciento en la rata. La fluorescencia específica atribuida a la dopamina está localizada en las células glómicas o epitelioides. Existe noradrenalina en menor proporción, en parte contenida en algunas células glómicas y en parte (50 por ciento aproximadamente) en fibras simpáticas que llegan por los nervios ganglio-glomerulares para inervar los vasos. La simpatectomía hace desaparecer los trayectos nerviosos fluorescentes, persistiendo la fluorescencia de las células glómicas.

(*b*) Estudios bioquímicos e inmunohistoquímicos muestran la presencia de tirosina-hidroxilasa y dopa-decarboxilasa (decarboxilasa de aminoácidos aromáticos) en las células glómicas; algunas células contienen también dopamina-β-hidroxilasa (dopamina-β-mono-oxigenasa); no se ha encontrado fenil-etanolamina-*N*-metil-transferasa. Los nervios ganglio-glomerulares son ricos en dopamina-β-hidroxilasa, enzima que sirve como trazador de neuronas noradrenérgicas. Las células glómicas serían entonces capaces de sintetizar dopamina y algunas podrían llegar a sintetizar noradrenalina, mientras la inervación simpática de los vasos sería noradrenérgica.

(*c*) No se ha investigado aún la presencia de catecol-orto-metil-transferasa (COMT) y existen observaciones dispares sobre la presencia de mono-amino-oxidasa (MAO) en el cuerpo carotídeo. No obstante, tratándose de sinapsis catecolaminérgicas, la reincorporación (o recaptación) importa más que la degradación enzimática como mecanismo de inactivación del transmisor.

(*d*) El cuerpo carotídeo contiene dihidroxi-fenil-alanina (DOPA) e incorpora activamente DOPA radiactiva. Los estudios autorradiográficos muestran que el trazador se deposita en los granos densos de halo claro.



FACTORES QUE ALTERAN el equilibrio iónico entre las células glómicas y su ambiente. El potencial de membrana y la resistencia de las células glómicas no son afectadas significativamente por cambios en la concentración externa de iones potasio. La remoción del ion sodio externo produce hiperpolarización y aumento de la resistencia, mientras que la reducción del ion cloro externo produce despolarización y caída de la resistencia. La hipotermia produce una reducción del potencial de membrana y de la resistencia de estas células; la hipertermia tiene el efecto contrario. La ilustración recoge acciones de estimulantes naturales y monoaminas endógenas en el cuerpo carotídeo. Hay un registro intracelular de célula glómica y extracelular de fibra nerviosa.

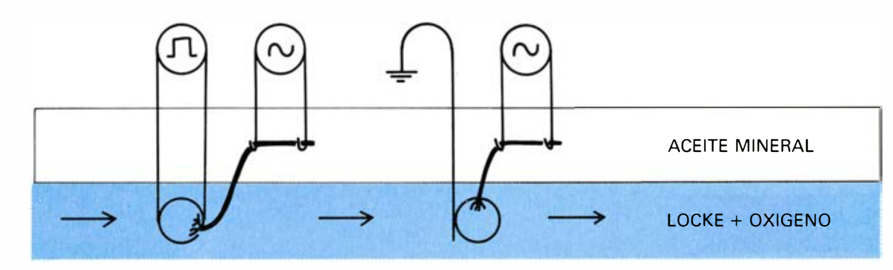
(*e*) La reserpinización aguda o crónica (que debiera reducir drásticamente los depósitos de catecolaminas) no modifica sustancialmente la respuesta quimiorreceptora a la hipoxia o la asfixia. Por otra parte, la 6-hidroxi-dopamina –que se incorpora a los granos densos de halo claro de las células glómicas– no produce degeneración de ellas ni altera las respuestas quimiosensoriales a asfixia, cianuro o dopamina.

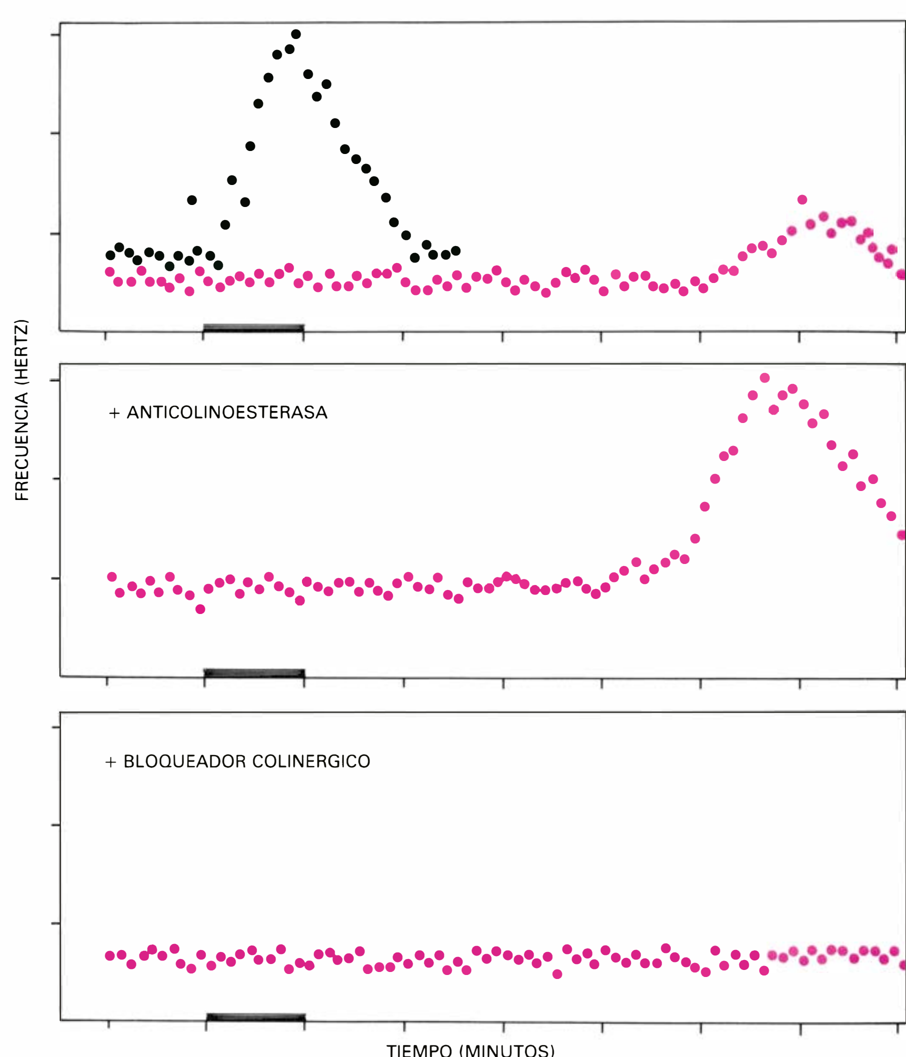
(*f*) Nos encontramos aquí con un problema esencial. La noradrenalina no tiene efecto quimiosensorial directo, sino sólo a través de efectos vasculares; sólo podría participar entonces en el control simpático de la circulación local del cuerpo carotídeo. Por otra parte, el efecto predominante de la dopamina en el hombre, el gato, el perro y la rata es una inhibición quimiosensorial. Por tanto, la dopamina no puede ser el mediador sináptico de la excitación quimiosensorial; no obstante, como veremos más adelante, podría desempeñar una función modulatoria de la quimiorrecepción. La situación puede ser diferente en el caso del conejo, en que la dopamina produce excitación.

(*g*) No disponemos de datos al respecto.

(*h*) Los bloqueadores α-adrenérgicos (dibenamina y fenoxibenzamina) y los β-adrenérgicos (propanolol y dicloroisoproterenol) no alteran la quimiorrecepción. Como ya vimos, las fenotiazinas y butirofenonas se comportan como bloqueadores dopaminérgicos en cuerpo carotídeo; ello –lejos de bloquear la quimiorrecepción– la potenciaría, habiendo podido comprobar un aumento de la actividad quimiosensorial luego de la administración de estas drogas.

(*i*) Los trabajos de Costa, Hanbauer y Hellström, en el Instituto Nacional de Salud, Washington, muestran que el contenido de dopamina del cuerpo carotídeo de la rata se reduce en relación a la severidad y duración de la hipoxia hipoxémica a que son sometidas. En el conejo también, Fidone, González y Yoshizaki han demostrado liberación de dopamina ante estos estímulos.



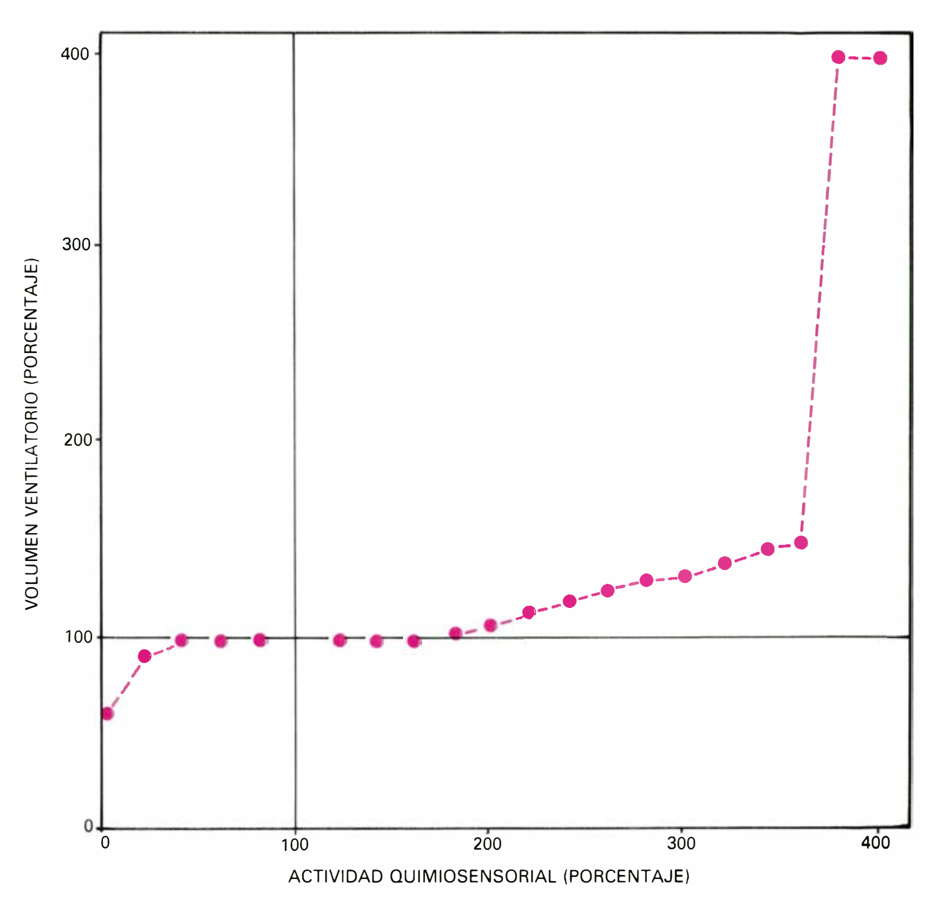


SI SE HACE PASAR CORRIENTE ELECTRICA a través de un cuerpo carotídeo superfundido in vitro, se registra de su nervio un aumento de frecuencia de descarga de los impulsos quimiosensoriales que no se establece de inmediato y se prolonga mucho más allá del término del estímulo. Ahora bien, si el líquido que baña este cuerpo se aprovecha para superfundir un segundo cuerpo carotídeo (*caso representado arriba*) se observa que tardíamente se produce un discreto aumento de la frecuencia de impulsos quimiosensoriales recogidos del nervio del segundo cuerpo carotídeo. Se trata, pues, de una superfusión de dos cuerpos carotídeos en serie, donde se ha verificado una estimulación con corriente eléctrica continua (durante un minuto) del primer cuerpo carotídeo, cuya descarga quimiosensorial se registra del nervio (*puntos negros*). El registro de actividad quimiosensorial de nervio del segundo cuerpo carotídeo, junto al que descansa la punta del electrodo de tierra, se dibuja en color rojo en las tres condiciones.

En conclusión, la estimulación natural del cuerpo carotídeo (al menos la hipóxica) produce liberación de dopamina. Pero la dopamina exógena produce inhibición (o excitación) de la descarga quimiosensorial. Por tanto, la dopamina endógena liberada modificaría la excitación quimiosensorial provocada por la hipoxia. A ello nos hemos referido al usar el término modulación.

El concepto de modulación es relativamente reciente y está en proceso de definición. Varios de los criterios de transmisión sináptica son aplicables a la modulación, por lo que parece necesario establecer una distinción entre mediador y modulador sináptico. El primero sería aquella sustancia sin cuya participación se detiene la transferencia de información fugaz del elemento presináptico al postsináptico. En cambio, el modulador no aparece como indispensable para la transmisión sináptica fugaz, pero puede ser un elemento determinante en la ganancia o ajuste fino del proceso.

En el Laboratorio de Neurobiología de la Universidad Católica de Chile, hemos estudiado la interacción entre dopamina y estimulantes del cuerpo carotídeo en el gato. Analizando curvas



RELACION ENTRE LA ACTIVIDAD QUIMIOSENSORIAL registrada de un nervio carotídeo seccionado del gato y los cambios ventilatorios reflejos mediados por el otro nervio carotídeo intacto (y los nervios aórticos). (Gráfico basado en observaciones precedentes de Zapata y Zuazo, 1980, publicadas en *Respir. Physiol. 40*: 79-92; y Serani y Zapata, 1981, en *Arch. Int. Pharmacodyn. Ther.*, 252, 284-297.)

dosis-respuesta al cianuro, se ha podido observar que la respuesta quimiosensorial máxima no es deprimida sino potenciada por la dopamina, mientras que las respuestas moderadas al cianuro son atenuadas. Con ello se descarta una simple sumación algebraica de efectos contrarios y se consolida un papel modulatorio para la dopamina.

Hay que agregar informaciones recientes sobre presencia en cuerpo carotídeo de varios polipéptidos: sustancia P, encefalina y polipéptido intestinal vasoactivo. ¿Participan estas sustancias en la regulación fisiológica de la actividad quimiosensorial? Es una interesante línea de investigación, pero la sola presencia de estas sustancias viene a confirmar la proposición de Fujita y Kobayashi de que estas células pertenecen al grupo de paraneuronas, células recepto-secretoras que almacenan hormonas peptídicas junto a transmisores catecolamínicos, concepto previamente esbozado por Pearse al describir un grupo de células que incorporan y decarboxilan a los precursores de aminas (APUD, del inglés "amine precursor uptake and decarboxylation"), donde se incluyen entre otras las células cromafines de la médula adrenal, las células pequeñas e intensamente fluorescentes de los ganglios autonómicos, las células endocrinas hipofisiarias, epifisiarias y gastro-intestinales, etcétera.

Existe abundante literatura sobre la importancia de los quimiorreceptores periféricos en la regulación de la ventilación pulmonar. En los últimos años estos estudios se han hecho extensivos a los seres humanos. Los tumores de cuerpos carotídeos –llamados quimiodectomas en reconocimiento a su papel quimiorreceptor– aunque relativamente raros se están reconociendo más tempranamente, pudiendo realizarse en dichos pacientes estudios pre- y post-operatorios. El mayor número de excisiones unilaterales y bilaterales de los cuerpos carotídeos se ha practicado en Japón y en los Estados Unidos en casos de asma bronquial resistente a otros tratamientos. También existen casos de endarterectomías carotídeas (para aliviar trombosis) que resultaron accidentalmente en quimiodenervación carotídea. Los estudios en pacientes con quimiodenervación carotídea han confirmado el papel preponderante del cuerpo carotídeo en la iniciación de la respuesta hiperventilatoria a la hipoxia, como también en la fase II (humoral) de tal respuesta al ejercicio.

Por otra parte, ha interesado cuantificar el grado de "comando" o influjo que ejercen los cuerpos carotídeos sobre el centro respiratorio en distintas circunstancias. La prueba más utilizada para ello fue elaborada por Pierre Dejours, actualmente en Estrasburgo, y consiste en la inhalación de dos bocanadas de O_2 puro para deprimir en forma abrupta los quimiorreceptores periféricos; la disminución ventilatoria que se produce con esta prueba sirve para medir la acción refleja excitadora que ejercen los cuerpos carotídeos sobre el centro respiratorio. Es posible que la hipoventilación provocada por la administración intravenosa de dopamina llegue a constituirse en otra prueba con este mismo objeto.

En el gato, existe la posibilidad de estudiar la relación entre la actividad quimiosensorial registrada de un nervio carotídeo y el efecto ventilatorio de la actividad quimiosensorial mediada por el nervio carotídeo contralateral intacto. La ilustración de la izquierda esquematiza dicha relación. La reducción de la frecuencia quimiosensorial de hasta 50 por ciento de la basal (mediante inyecciones de dopamina) no modifica la ventilación, pero al reducir transitoriamente a 0 la frecuencia quimiosensorial, la ventilación se reduce hasta un 60 por ciento. Por el contrario, el aumento de la frecuencia quimiosensorial (por inyecciones de cianuro) de hasta 200 por ciento de la basal no altera la ventilación; mayores incrementos quimiosensoriales producen incrementos proporcionales de la ventilación, hasta que se alcanza determinado umbral que se traduce en la aparición de respiración jadeante.

La exposición anterior corresponde a una visión parcial de los problemas planteados por el estudio de la fisiología del cuerpo carotídeo, analizado como modelo de receptor complejo y donde se le ha dado especial cabida a aquellos interrogantes que inquietan a los autores de este trabajo.

El cuerpo carotídeo atrae gran interés en estos últimos años. Son varios los laboratorios en distintos países donde se está investigando activamente al respecto. Este interés puede quizás expresarse en que habiéndose organizado un primer simposio internacional sobre Quimiorreceptores Arteriales en Oxford, Inglaterra, en 1966, rápidamente lo siguieron otros en Chicago, Estados Unidos (1971), Bristol, Inglaterra (1973), Srinagar, India (1974), Dortmund, Alemania Federal (1976), Dallas, Estados Unidos (1979) y Valladolid, España (1979).

Temas metamágicos

De la memez numérica, o cómo el anumeralismo engendra iguales peligros que el analfabetismo

Douglas R. Hofstadter

La profesora Grandnumskaia, afamada cosmofísica, disertaba acerca de la futura evolución del Universo. Acababa de explicar que, según sus cálculos, al cabo de un billón de años la Tierra terminaría cayendo en el Sol, encontrando allí ígnea muerte. Del fondo de la sala se alzó, trémula, una voz. "Discúlpeme Vd., profesora, pero ¿cu-cuánto ha dicho que falta?" Grand numskaia respondió tranquilamente: "Alrededor de un billón de años". Se oyó un suspiro de alivio. "¡Gracias a Dios! ¡Por un instante me pareció haber oído un *millón*!"

El presidente Kennedy solía referir una anécdota relativa a un famoso soldado francés, el mariscal Lyautey. Un día, el mariscal le encargó a su jardinero que a la mañana siguiente plantase en el jardín una hilera de árboles de cierta rara variedad. Aunque el jardinero accedió gustoso, previno al mariscal de que ese tipo de árboles tardaba un siglo en alcanzar completo desarrollo. "En tal caso", replicó Lyautey, "plántelos esta misma tarde".

En estas dos anécdotas, un suceso futuro muy distante se compara de forma sorprendente con otro mucho más a nuestro alcance. En la segunda historia pensamos para nosotros: con un siglo por delante, ¿qué importancia puede tener un día? De todas formas, nos encanta el agudo sentido de apremio del mariscal. Cada día cuenta, parece decirnos, especialmente cuando hay miles de ellos. Siempre he encontrado esta anécdota de mi gusto; la otra se me hace más chistosa. La idea de que alguien pueda sentir que números tan enormes le afectan directamente, de que el fin del mundo pueda sentirse como algo más cercano si faltasen solamente un millón de años y no un millón de millones... ¡es ridículo! ¿Quién podría reaccionar tan visceralmente al captar la diferencia entre números tan enormes?

He sospechado siempre que pocas personas alcanzan a ver la diferencia entre un millón y mil millones, o entre mil millones y un billón. Desde luego, sí la perciben lo suficientemente como para verle la gracia al chistecito de la Tierra y el Sol; pero en qué consiste la diferencia, eso ya es harina de otro costal. He oído a un locutor de radio declarar "La sequía ha costado entre 900.000 y mil millones de dólares a la agricultura californiana". Cosas así me preocupan mucho. En sociedades como las nuestras, donde las cantidades grandes son moneda corriente, no podemos permitirnos un desconocimiento tan pasmoso de su significado. ¿O es que verdaderamente padecemos de memez numérica? ¿Estamos haciéndonos cada vez más insensibles a los números, que cada vez son más grandes?

¿Qué piensa la gente de que en algunas de nuestras ciudades más densamente pobladas el precio de un piso bien situado pueda alcanzar los 20 millones de pesetas? ¿Y de los anuncios de ciertas instituciones de ahorro, que proponen sólo inversiones millonarias? ¿Acaso *todo el mundo* puede ser multimillonario? ¿Debemos *esperar* que una casa le cueste a uno la quinta parte de CIEN millones? ¿Qué ha sido de las relumbrantes connotaciones de la palabra "multimillonario"?

Hace años fui instructor de un pequeño grupo de alumnos principiantes en ciencias físicas. Dábamos clase en la planta 13 del Hunter College, en Nueva York. Desde las ventanas teníamos una espléndida panorámica de los rascacielos de Manhattan. En una de las primeras clases quise enseñar a mis alumnos las ideas de valor aproximado y de cifras significativas, y para ello les pedí que estimasen la altura del Empire State Building. En un grupo de diez estudiantes, ni uno sólo llegó a un valor comprendido entre la mitad y el doble del correcto (448,7 metros, incluida la antena de televisión). La mayoría de las respuestas estuvieron entre 100 y 150 metros. Uno dijo 15 metros –verdaderamente una subestimación que nos deja atónitos. Otro dijo que kilómetro y medio. Resultó que este último sí había calculado la solución, suponiendo que cada piso pudiera tener unos 15 metros, y que el número de pisos sería de unos 100. Donde una persona opinó que cada planta pudiera tener unos 15 metros de altura, otra pensó que un edificio de nada menos que 102 pisos sólo alcanzaría ese total. Tan estupefaciente episodio ejerció en mí profundo efecto.

Está hoy *à la page* lamentar el deplorable y pasmoso analfabetismo de esta generación, incapaz, se dice, de escribir conforme a la gramática. Mas, ¿qué decir del sobrecogedor "anumeralismo" de casi todo el mundo, jóvenes y viejos, cuando se trata de interpretar y comprender las cifras que rigen sus vidas?

El mundo es grande, qué duda cabe. Hay en él mucha gente, muchas necesidades; y todo ello produce en conjunto cierta dificultad de comprensión. Esta dificultad no puede ser excusa para negarse a comprender, e incluso a referirse, a números cuya finalidad es recoger en unos pocos símbolos ciertos aspectos sobresalientes de esas enormes realidades. Verosímilmente, no serán los lectores de este artículo quienes deban preocuparme en este respecto. Es, sin embargo, cierto que cada uno de mis lectores conoce a muchas personas que se encuentran incómodas al enfrentarse a los grandes números que aparecen en el PNB, en los presupuestos del Estado, en los balances de sociedades bancarias y demás. A quienes los números terminados en "-illones" les dejan la mente en blanco, todos los números grandes les resultan iguales, y variaciones verdaderamente explosivas en las cantidades manejadas no les inmutan. Semejante incapacidad para manejar números grandes es, a todas luces, socialmente perniciosa. Induce a la gente a dejar de lado cuestiones fundamentales porque, sencillamente, las encuentran ininteligibles. Por tanto, todo cuanto pueda hacerse para corregir este "anumeralismo" galopante bien valdrá la pena.

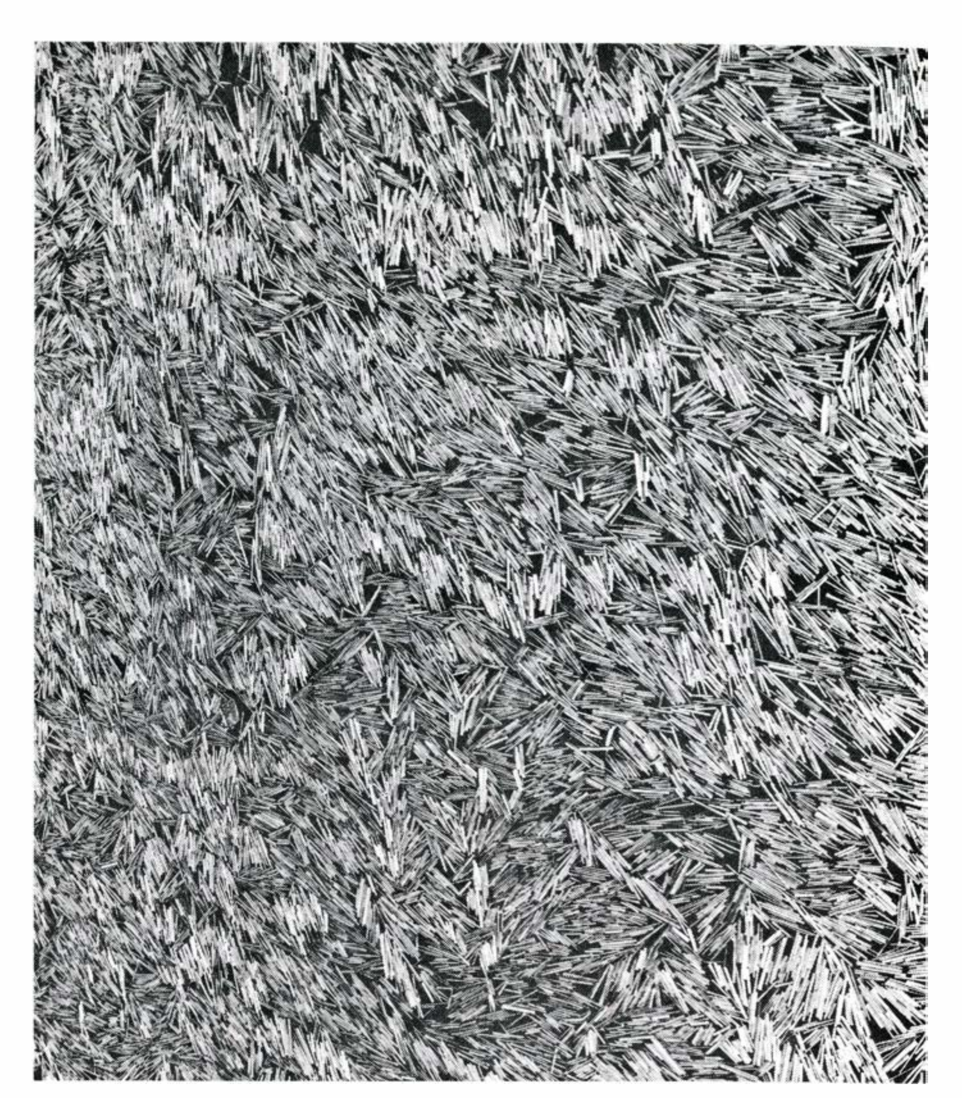
Como ayuda para adquirir un poco de sentido común al manejar números, voy a consentirme una pequeña orgía de preguntas y respuestas. Vamos allá. ¿Cuántas letras cree usted que contiene esta revista? No las calcule, dígalo a ojo. ¿Ha contestado un millón? Esa cifra tiene seis ceros. Si así ha sido, su estimación ha sido muy razonable. Si no, ¿se pasó usted por alto o por bajo? ¿Le parece ahora que su estimación ha sido descabellada? ¿Qué clase de pistas, de indicios intuitivos, le sugieren que un millón pueda ser cifra más apropiada que, pongamos por caso, unos cuantos miles, o que mil millones? Más vale que ahora calculemos.

Pongamos que hay cien páginas de texto compacto (seguramente estemos sobrestimando). ¿Cuántas palabras por página 500 o 5000? Más bien nos parece que una cifra intermedia. Pongamos unas 2000. ¿A razón de cuantas letras por palabra? Ah, pues unas seis. Tendremos entonces 100 × 2000 × 6, o sea, algo más de un millón. Quizás sea un valor demasiado alto, pero sin duda no está desencaminado. Mas, cómo hubiéramos podido intuir este número por adelantado?

Hemos, pues, de encarar una elección. ¿Cuál es la más verosímil de las 12 posibilidades siguientes: (*a*) 10; (*b*) 100; (*c*) 1000; (*d*) 10.000; (*e*) 100.000; (*f*) 1.000.000; (*g*) 10.000.000; (*h*) 100.000.000; (*i*) 1.000.000.000; (*j*) 10.000.000.000; (*k*) 100.000.000.000; (*l*) 1.000.000.000.000? En principio la dificultad de una estimación semejante es más o menos igual a la de echar una ojeada a una habitación y estimar si hay dos sillas, o siete, o 15. Lo único que pasa es que aquí queremos saber el número de ceros de una cantidad, es decir, que estamos atendiendo al logaritmo (en base 10) de tal cantidad. Si de una mirada somos capaces de "captar" cuántos asientos hay en una habitación, por qué no ser igualmente capaces de "percibir" el número de ceros de una cifra?

Claro está, hay diferencia. Una cosa es echar una ojeada a una billonada, como 10000000000000 y tener la impresión intuitiva, sin contarlos, de que hay alrededor de 12 ceros, desde luego más de diez y menos de 15, y otra muy distinta tener la impresión, al contemplar una fotografía aérea de una masa de troncos cortados, de que la representación decimal del número de troncos deberá tener de tres a cinco ceros, es decir, que 1000 es una estimación decididamente baja, y que 100.000 es exageradamente alta. Esta capacidad ya es una forma de abstracción, un peldaño superior al de la percepción numérica habitual. Este primer nivel de abstracción, empero, no debiera resultar demasiado difícil de alcanzar.

El secreto, desde luego, es la práctica. Es preciso habituarse a la idea de que 10 ceros son muchos ceros para que un número alcance a tenerlos, de que los números de cinco cifras son cantidades francamente grandes, y de que los de tres representan valores intuitivamente comprensibles. Con tres ceros basta ya para dar cuenta del número de alumnos de un instituto de bachillerato, 1000, salvo un factor o divisor igual a 3. (En los números de unos cuantos ceros hemos de estar dispuestos a tolerar siempre un coeficiente tres hacia arriba o abajo, siempre y cuando nos interese una estimación, y no el valor exacto de un caso concreto.) Cinco ceros expresa el número de personal de una comarca rural importante, unas 100.000 almas o así. Seis ceros, es decir, millones, representan grandes ciudades –Valencia, o Bilbao, o Marsella o Brasilia, o Dar es Salaam. Siete ceros denotan ya urbes enormes, Shanghai, o Ciudad de México, o Seúl, o París, o Nueva York. ¿Cuántas cree usted que serán las ciudades del mundo que alcancen o rebasen el millón de habitantes? Es en casos así donde la práctica sirve de ayuda.



¿Cuántos troncos vemos en esta fotografía aérea?

¿Cuántos árboles han de talarse semanalmente para producir la edición dominical del *New York Times*? Pongamos que se impriman un par de millones de ejemplares, cada uno de unos dos kilos de peso. Tenemos así 4 millones de kilos de papel, o sea, 4000 toneladas. Si cada árbol diera para una tonelada de papel harían falta 4000 árboles. No soy muy buen cubero, pero no creo que nos despistemos mucho admitiendo una tonelada por árbol. En el peor de los casos pudieran ser 100 kilos de papel por árbol, equivalentes a 40.000 árboles pequeños. La fotografía de esta página muestra una masa de 7500 a 15.000, a mi juicio. Por consiguiente, admitiendo 100 kilos de papel por árbol, los troncos de la fotografía representan menos de la mitad de un dominical del *Times*. Podríamos ahora tratar de estimar el número de árboles talados mensualmente para editar todas las revistas, periódicos y libros, pero dejaré este ejercicio a cuidado del lector.

¿Cuántos cigarrillos se fuman anualmente en los Estados Unidos? ¿Cuántos ceros harán falta? Una buena docenita, pues la cifra es del orden del billón. No es difícil de calcular. Pongamos que la mitad de la población fume cigarrillos: unos 100 millones. Ya sé que esta cifra de fumadores es excesiva; luego la compensaremos reduciendo algo. Cada fumador consume ¿cuánto? ¿Una cajetilla diaria? Sea. Tenemos así 20 cigarrillos por 100 millones: 2000 millones de cigarrillos diariamente. Hay 365 días en el año. Pongamos sólo 250,

puesto que prometí rebajar en algún sitio. 250 por 2000 millones nos da 500.000 millones –medio billón. Por poco le damos en toda la nariz. El dato que tengo, de hace pocos años, dice 545.000 millones.

Son números de este tamaño los que se manejan al decir que el presupuesto de defensa estadounidense de los próximos cuatro años rondará los 750.000 millones de dólares. Un ordenador personal (del tipo que al autor le gustaría poseer), con todos sus periféricos, viene a costar unos 75.000 dólares. Con 750 mil millones para gastar, podríamos entregarle uno así a cada habitante de Nueva York, es decir, podríamos adquirir unos 10 millones de ellos. Podríamos también darle 1 millón de dólares a cada empadronado en San Francisco, y todavía nos quedaría lo suficiente para regalarle una bicicleta a cada chino. No hay forma de imaginar todos los buenos fines a que pueden dedicarse 750.000 millones de dólares. Pero se invertirán en municiones, tanques, aviones de combate, maniobras militares, sistemas de misiles, combustible de aviación y demás. Es una forma interesante de gastar 750.000 millones, aunque a mí se me ocurren otras mejores.

Pensemos en números grandes de otros tipos. ¿Sabía usted que la retina humana contiene alrededor de 60 millones de células, cada una de las cuales responde a un tipo de estímulo particular? Todas ellas envían sus señales al cerebro, que hoy se supone formado por unos 100.000 millones de neuronas, o sea, células nerviosas. El número de células gliales, células menores con funciones de soporte y sostén, contenidas en el cerebro, se estima en unas 10 veces mayor. Pueden parecernos números grandes, pero nuestro cuerpo está integrado por unos 60 o 70 billones de células. Cada una de ellas contiene millones de componentes trabajando en colaboración. Tomemos por ejemplo la hemoglobina, proteína encargada del transporte de oxígeno en la corriente sanguínea. Cada uno de nosotros tiene unos 6000 millones de billones, o sea, unos 6000 trillones, de moléculas de hemoglobina, de las cuales ¡unos 400 billones se destruyen y otras tantas se fabrican cada segundo!

El número de moléculas de hemoglobina de nuestros organismos es de 6×10^{21}. Resulta curioso que durante los dos últimos años mucha gente se haya familiarizado con un número casi igual de grande, a saber, el número de posibles configuraciones distintas del cubo de Rubik. Este número –llamémoslo constante de Rubik– es aproximadamente $4{,}3 \times 10^{19}$. Más aún, si se trabaja con el "supergrupo" de Rubik, donde también importan las orientaciones de las caras centrales, entonces la constante de Rubik es ¡alrededor de 9×10^{22}!

El número de hemoglobinas y la superconstante de Rubik son números *auténticamente* grandes. ¿Qué les parece si nos fijamos en números más pequeños, aunque sea por pisar suelo sólo un momento? Muy bien, ¿cuántas personas cree usted que están descendiendo a tierra en paracaídas en este momento (que es de presumir sea un momento perfectamente típico)? ¿Cuántas palabras conoce usted de su idioma materno? ¿Cuántos asesinatos se cometen todos los años en el Condado de Los Angeles? ¿Y en Japón? Estos dos últimos números nos chocan muchísimo al compararlos: en el Condado de Los Angeles, unos 2000. En todo el Japón, unos 900.

Y hablando de muertes violentas, he aquí unas que todos procuramos barrer debajo de la alfombra, a lo que parece. Sólo en Norteamérica fallecen en accidente de circulación unas 50.000 personas todos los años. ¿Podemos imaginar nuestra reacción si alguien nos dijera de pronto: "¡Eh, fíjese! Se me acaba de ocurrir un invento verdaderamente majo. Tiene un solo defecto, que cada 10 o 12 años borraría del mapa una ciudad más o menos como San Francisco. ¡Pero, no se vaya! Les prometo que a todos los demás va a encantarles. ¡De verdad!" Ahora bien, estos datos son sensiblemente exactos para los coches. Pocas veces oímos decir a la gente que los coches deban prohibirse. De alguna manera estamos dispuestos a aceptar el coste de 50.000 vidas anuales sin mayores preocupaciones.

Quizá convenga tratar la cuestión un poco más a la ligera. ¿Y qué hay más ligero que la luz? La luz está compuesta de fotones. ¿Cuántos fotones emite cada segundo una bombilla de 100 watt? Alrededor de 10^{20} –otro de los números grandes. ¿Es este número mayor o menor que el número de granos de arena de una playa? ¿De qué playa? Digamos que una playa de 1 kilómetro de largo, 30 metros de anchura y con una capa de arena de un metro y medio de espesor. ¿A cuánto estima usted que ascenderá ese número? Bien, calcúlelo ahora. ¿Y qué me dice del número de gotas de agua del Océano Atlántico? Pruebe entonces con el número de peces del océano. ¿Quiénes serán más, los peces del mar o las hormigas en tierra? ¿El número de átomos de una hoja de hierba o el número de hojas de hierba de toda la Tierra? ¿Las hojas de hierba o los insectos? ¿El número de hojas de un roble corriente o el de cabellos de una cabeza humana? ¿Cuántas gotas de lluvia caen en su ciudad en un segundo durante un fuerte aguacero?

¿Cuántas reproducciones de la Mona Lisa se habrán impreso hasta hoy? Probemos juntos a calcularlo. Seguramente aparezca unas cuantas docenas de veces al año en las revistas americanas. Pongamos que cada revista lance unos 100.000 ejemplares. Tendremos así unas cuantas docenas de millones de copias anuales, contadas solamente las revistas; aparte debemos considerar libros y otras publicaciones. Tal vez debiéramos duplicar o triplicar esta cifra, sólo en los Estados Unidos. Teniendo en cuenta los restantes países, podremos multiplicar nuevamente por tres o cuatro. Ahora hemos alcanzado ya cifras próximas a los 100 millones anuales. Supongamos que tal cantidad haya sido válida para todos los años, de un siglo a esta parte. ¡Tendríamos entonces unos 10.000 millones de reproducciones de la Mona Lisa! Sin duda hemos cometido errores por el camino, pero multiplicando o dividiendo por 10, éste será el orden de magnitud del número.

¿Multiplicar o dividir por 10? Hace un momento estaba yo diciendo que un factor tres podría tolerarse; ahora estoy concediéndome más de *dos* factores tres, es decir, todo un orden de magnitud, un factor 10. Bueno, pero es que ahora estamos manejando números mucho mayores (10^{10} en lugar de 10^5); por ello nos es permisible el factor 10. De paso, podemos dar por sentada una regla de buen sentido, a saber, admitir un coeficiente de tolerancia igual a 3 por cada factor 100.000 que contenga el número. Significa esto que podemos desviarnos del valor correcto en un factor 10 –un orden de magnitud– cuando se alcancen cifras del orden de los 10.000 millones, en un factor 100 o así (dos órdenes de magnitud) al subir hasta 10^{20} (cien trillones), un número aproximadamente 2,5 veces mayor que la constante de Rubik.

¿Y por qué debemos darnos por contentos con una estimación que puede ser sólo el 1 por ciento del valor real, o por otra parte, tal vez cien veces mayor? Bien, porque al tomar el logaritmo decimal del número –el número de ceros– si decimos 18 y la respuesta correcta es 20, ¡sólo nos equivocamos en el 10 por ciento!

¿Y qué derecho tenemos a prescindir tan altivamente de la propia cantidad, y

fijar la atención solamente en su logaritmo, en su "orden de magnitud"? Sencillamente, que cuando los números alcanzan tales tamaños ya no hay otra solución. La realidad perceptual comienza a transmutarse. Somos *incapaces* de visualizar la cantidad verdadera; la cifra –la hilera de guarismos– pasa a ocupar su lugar, y nuestra realidad perceptual se convierte en la del número de ceros. ¿En qué momento se produce esta mutación? Cuando el ojo de la imaginación ya no logra ver una colección del orden de magnitud apropiado. Para mí, esta confusión perceptual comienza en torno a 10^4 –el tamaño del revoltijo de troncos que ha captado la fotografía.

Hay otras formas de percibir 10^4, como por ejemplo, el número de botes de conserva que ocuparían una estantería de 15 metros de un supermercado. Cuando los números son mucho mayores que éste, a mí me es sencillamente imposible visualizarlos. El número de azulejos que recubren el Lincoln Tunnel, entre Manhattan y New Jersey, es tan enorme que a duras penas logro representármelo. (Debe ser del orden de un millón.) Sea como fuere, en el intervalo de 10^4 a 10^5 mi capacidad de visualización comienza a desdibujarse y a ser sustituida por otra realidad de segundo orden, indirecta, que es el número de dígitos (o hasta cierto punto, los nombres de los números, como millón, billón, trillón...). Por qué se produce el desenfoque a este tamaño y no por ejemplo en 200 unidades o en 10 millones, seguramente tenga que ver con la evolución y con el papel que la percepción de estas vastas realidades represente en la capacidad de supervivencia. Es una cuestión filosófica fascinante, pero que no cabe responder aquí.

En cualquier caso, he aquí una regla de buen sentido: La estimación debería encontrarse dentro de un margen del 10 por ciento de la respuesta correcta, pero sólo hace falta que tal sea cierto *a nivel de nuestra realidad perceptual*. Por consiguiente, estaremos bien disculpados si estimamos que el Cubo de Rubik admite 10^{18} posiciones, porque 18 está francamente próximo a 191/2, que es casi exactamente igual al número de cifras. (El factor 4,3 cuenta por un poco más de media cifra, pues cada factor 10 aporta una cifra completa y, cada factor 3,16, la raíz cuadrada de 10, contribuye con medio dígito.)

Tal vez, si tuviéramos que empezar a manejar números con millones o billones de cifras, los propios números (que en nuestra notación serían interminables hileras de guarismos) dejarían de ser visualizables y nuestra realidad perceptual tendría que saltar a niveles de abastracción superiores: al número que cuenta los dígitos del número que cuenta los dígitos del número que expresa cuántos son los objetos a que nos referimos. No es preciso decir que la realidad perceptual de tercer orden es abstracta en grado sumo. Además, muy raramente se presenta, ni aun en matemáticas. Empero, podemos imaginar cómo rebasarla con mucho. Las realidades perceptuales de cuarto y quinto orden rápidamente abrirían paso, en nuestra imaginación puramente abstracta, a realidades de órdenes décimo, centésimo, millonésimo.

Evidentemente, llegados a este punto habríamos perdido la cuenta del número exacto de niveles rebasados, y tendríamos que contentarnos con la mera *estimación* de tal número (¡con tolerancias del 10 por ciento en más o en menos, lógicamente!). "Pues... a mi juicio, en este problema nos hemos elevado hasta un par de millones de niveles de abstracción conceptual, más menos un par de centenas de millar", pudiera muy bien ser el comentario de doctas personas expertas en el manejo de cantidades tan inimaginablemente inimaginables. Ya vemos hacia dónde eso nos lleva: a referirnos a múltiples niveles de abstracción desde múltiples niveles de abstracción. Si fuera preciso prolongar esta discusión un lelosegundo más íbamos a vernos encallados en mitad de las teorías de funciones recursivas y complejidad algorítmica, y eso *sí* es demasiado abstracto. Por lo tanto, abandonamos la cuestión aquí mismo.

El cálculo de cifras del famoso número π, aunque mucho más tangible, también guarda relación con estas ideas. ¿Cuántos decimales se han calculado hasta ahora, con auxilio de máquinas? La respuesta es (que yo tenga noticia) un millón. La tarea se llevó a cabo en Francia, y el millón de guarismos llena todo un libro. De ese millón, ¿cuántas han podido encomendarse a una memoria humana? La respuesta provoca incredulidad: 20.000, según la última edición del *Guinness Book of World Records*. Yo mismo llegué a aprenderme, siendo estudiante un poco chalado, 380 cifras. He conocido más tarde otras personas que me han superado. Casi todos nosotros hemos olvidado la mayor parte de las cifras que supimos, pero todavía teníamos seguridad en las 100 primeras, y en alguna ocasión, las hemos recitado juntos, al unísono. Algo esotérico el capricho.

¿Qué pensaría el lector si alguien le dijera que una persona ha llegado a aprenderse de memoria el libro completo de cifras de π? Estoy seguro que nadie prestará oído a semejante patraña. En cierta ocasión oí decir que Jerry Lucas, el famoso fenómeno del baloncesto y no menos conocido memorión, había afirmado saberse de memoria la guía telefónica de Manhattan. Tenemos aquí buen ejemplo de cómo el anumeralismo amamanta la credulidad. ¿Puede el lector imaginar lo que tal proeza supondría? A mi juicio excede al menos en dos órdenes de magnitud lo más que podríamos creer. Aprender de memoria una página de la guía ya me parece de una dificultad fabulosa. Aprender de memoria 10 páginas sin duda roza el límite de lo creíble. Incidentalmente, la completa memorización de la Biblia (cosa que en ocasiones he oído proclamar) me parece equivalente a la memorización de 10 páginas de la guía telefónica. Ser capaz de lograrlo con 1500 páginas de números telefónicos, nombres y direcciones está, literalmente, más allá de lo posible.

Fijémonos en la escala con que se miden determinados fenómenos. Por ejemplo, el tono musical. Mirando el teclado de un piano vemos una escala lineal con la cual medir un tono audible cualquiera. Lo más natural es decir cosas como "Este la se encuentra nueve semitonos más alto que el do, y el do, siete semitonos más alto que ese fa, así que el la se encuentra 16 semitonos por encima del fa". Es una escala aditiva; lo que significa que a cada nota se le puede asignar un número, y la distancia de una a otra nota será la diferencia entre los números que les corresponden.

Por otra parte, desde el punto de vista acústico, resulta más conveniente describir los tonos por su frecuencia que por su situación sobre el teclado. La frecuencia es otro número asociado a cada una de las teclas. En lo más bajo, el la más grave vibra a razón de unas 27 veces por segundo, mientras que el do situado tres semitonos más arriba lo hace a 32. Pudiéramos sentirnos inclinados a pensar que para subir otros tres semitonos habrá que ir añadiendo siempre cinco ciclos por segundo. No es así. Lo que debe hacerse, en cambio, es *multiplicar* siempre por un valor muy cercano a 32/27.

Así pues, cuando hayamos subido una octava, que consta de 12 semitonos, la frecuencia del tono inicial ha sido multiplicada cuatro veces consecutivas por 32/27, es decir, por 2. En realidad, la cuarta potencia de 32/27 no es 2 exactamente, y puesto que la octava corresponde a una frecuencia exactamente doble, 32/27 tiene que ser tan sólo

una aproximación. Empero, no es esto lo esencial. Lo esencial es que las operaciones naturales para la comparación de frecuencias son la multiplicación y la división, mientras que las operaciones naturales con las notas del teclado son adición y sustracción. En el fondo, esto significa que los números asignados a las notas son logaritmos de frecuencias. He aquí un ejemplo donde resulta natural pensar logarítmicamente.

Se piensa logarítmicamente cuando sólo se perciben incrementos lineales al duplicarse la magnitud de la cosa percibida. Por ejemplo, ¿nunca se ha sentido maravillado de que marcando siete cifras sea posible poner en comunicación dos teléfonos cualesquiera del área metropolitana de grandes ciudades como Madrid o Barcelona, o Nueva York, donde residen 10 millones de personas? Supongamos que Nueva York duplicase su población. ¿Sería necesario añadir siete cifras más a cada número de teléfono, y tener que marcar 14 cifras para poder comunicar con estos 20 millones de personas? Desde luego que no. Añadir otros siete dígitos *multiplicaría* por 10 millones el número de posibilidades. Tanto es así, que basta anteponer únicamente tres cifras (dos para Barcelona, Madrid, Valencia y Vizcaya) para poder alcanzar desde cualquier punto de España cualquier otro teléfono del país. Así sucede porque cada nueva cifra multiplica por diez el número de teléfonos alcanzables. Tres dígitos más multiplicarán siempre por mil la red conectada al teléfono propio: tres órdenes de magnitud. Por consiguiente, la longitud de un número telefónico –la magnitud directamente percibida cuando protestamos de lo mucho que se tarda en marcar en las conferencias interurbanas– da una medida logarítmica del tamaño de la red de comunicaciones en la que estamos inmersos. Por esa razón vemos cuán ridículos son estos números enormemente largos, formados por 20 o 30 dígitos, que se utilizan para codificar personas o productos, cuando sin duda bastarían unos cuantos guarismos.

En cierta ocasión recibí una notificación pidiéndome que efectuara una transferencia a la cuenta número 60802-620-1-1-721000-421-01062 de un banco de Yugoslavia. Durante cierto tiempo esta cifra detentó mi record personal de números absurdos relativos a transacciones comerciales. Recientemente, sin embargo, al recibir los impresos de registro de mi automóvil encontré al pie la siguiente constante aclaratoria:

010101361218200301070014263117241512003603600030002.

Por si acaso, tras unos cuantos espacios en blanco, llevaba de apostilla 19283.

Un ejemplo donde se manifiesta el pensamiento logarítmico es en los nombres de los números. En Europa los números grandes cambian de nombre cada seis ceros. (En Estados Unidos lo hacen cada tres. Allí un billón son mil millones; un trillón, mil de sus billones, o sea, un millón de millones.) Tenemos así millón, billón, trillón... En cierto sentido, cada salto es de "igual tamaño". Así, un billón es tantas veces mayor que un millón cuantas un trillón lo es de un billón. Más directamente, un billón es a un millón como un trillón es a un billón. (Por otra parte, ¿debemos decir 10^{103} es a 10^{100} exactamente lo mismo que un millón es a un millar? Yo me siento inclinado a contestar que no. La respuesta es delicada a causa de los deslizamientos de la realidad perceptual, pues aquí no nos interesa tanto la aritmética cuanto la psicología de la percepción de números grandes.)

En cualquier caso, parece como si al llegar al trillón se nos agotasen los nombres de los números. En los diccionarios, qué duda cabe, figuran las denominaciones oficiales de números mayores, que nos son más o menos igual de conocidas que los nombres de los extintos dinosaurios: cuatrillón, vigintillón, brontosillón, triceratrillón,... Sencillamente no nos resultan familiares, pues murieron hace más o menos un dinosillón de años. Incluso "billón" suscita dificultades, al tener significados distintos en distintas culturas, como ya se ha mencionado. ¿Puede usted hacerse idea de las confusiones que se suscitarían si en el resto de Europa la palabra "cien" significara 1000? La verdad es que la imaginación de la gente repudia los números demasiado grandes. Es una lástima, de todas formas, que un trillón sea el mayor de los números con nombre corriente. ¿Qué ocurrirá cuando los presupuestos de defensa lo rebasen? ¿Llegaremos entonces a insensibilizarnos aún más?

También la escala logarítmica se adapta mejor para consignar los avances en cálculo automático. A lo largo de estos últimos decenios, el número de operaciones primitivas (como la adición y la multiplicación) que los ordenadores pueden ejecutar en un segundo ha venido decuplicándose cada siete años. Actualmente es del orden de 100 millones de operaciones por segundo; algo mayor en las máquinas más refinadas. Hacia 1975 esta cifra rondaba los 10 millones de operaciones por segundo. A fines de la década de los sesenta se consideraba que un millón de operaciones era una velocidad excepcional; a principios de esa década, se tenía como tal las 100.000, a mediados de los cincuenta, 10.000; a fines de los cuarenta, 1000, y a su comienzo, 100.

En efecto, en los primeros años cuarenta, Nicholas Fattu era director de un equipo de la Universidad de Minnesota que estaba trabajando en ciertos cálculos estadísticos para la Fuerza Aérea donde intervenían matrices grandes (de unos 60 por 60). Reunió en una sala a unos diez calculistas, y los proveyó a cada uno de una calculadora de sobremesa "Monroematic". Este equipo estuvo trabajando coordinadamente, con dedicación completa, durante 10 meses, ejecutando al mismo tiempo que el cálculo las comprobaciones recíprocas necesarias. Veinte años más tarde, el profesor Fattu ejecutó por curiosidad aquellos cálculos en una IBM 704, que invirtió veinte minutos en ello. Descubrió también que el equipo de cálculo había cometido dos errores intrascendentes. En nuestros días, la tarea completa podría realizarla en uno o dos segundos un gran ordenador.

Mas tampoco los ordenadores modernos son infinitamente veloces. La notable demostración del teorema de los cuatro colores llevada a cabo en la Universidad de Illinois hace algunos años exigió 1200 horas de ordenador. Traducido a días aún se nos hace más impresionante: nada menos que 50 días completos, a 24 horas cada uno. Si el ordenador estuvo ejecutando 20 millones de operaciones por segundo, la masa total de cálculos ascendería a unos 10^{14}, o sea 100 billones, de operaciones elementales, equivalentes a un par de centenares por cada cigarrillo fumado ese año en los Estados Unidos. ¡Uf!

Un ordenador capaz de ejecutar 1000 millones de operaciones por segundo ya llevaría buena marcha. Imaginemos un segundo descompuesto en tantos minúsculos instantes cuantos segundos componen 30 años. Así de pequeño es un nanosegundo, una milmillonésima de segundo. Para un ordenador tal, un segundo es toda una vida. Aunque, claro, ese ordenador es una tortuga comparado con los sucesos que se producen en los átomos que lo forman. Tomemos uno de ellos. Un electrón típico circunvala a un núcleo típico unas 10^{15} veces por segundo, es decir, un millón de veces en un nanosegundo. Visto con ojos de electrón, el ordenador corre tan lento como la miel en enero.

Porque el electrón tiene verdaderamente dos ojos con los que observar la

situación: además del período orbital tiene un período de rotación sobre sí mismo, de *espín*. Ahora bien, estrictamente hablando, "espín" no es más que una metáfora en la mecánica cuántica, así que cuanto sigue es conveniente tragarlo con una pizca de sal. De todas formas, si el electrón se imagina como una esfera giroscópica clásica (es decir, no mecánico-cuántica) puede calcularse su período de rotación a partir del momento angular de espín, que es conocido (y sensiblemente igual a la constante de Planck, unos 10^{-34} joule-segundo) y de su radio (que podríamos considerar igual a su longitud de onda Compton, alrededor de 10^{-10} centímetros). El período de espín resulta entonces del orden de 10^{-20} segundos. Dicho de otra forma, por cada vez que nuestro superveloz ordenador suma dos números, cada uno de los electrones de su interior ha ejecutado 100.000 millones de piruetas en torno a su eje. (Si hubiéramos tomado en cambio el radio "clásico" del electrón, éste hubiera dado 10^{24} rotaciones sobre sí mismo en un segundo ¡suficientes para marear a cualquiera! Como esta cifra viola tanto las limitaciones impuestas por la mecánica cuántica como la relativista, nos contentaremos con el primer valor.)

En el otro extremo de la escala tenemos la lenta y majestuosa revolución de nuestra galaxia, que requiere para dar su sereno giro unos 200 millones de años. Dentro del Sistema Solar, el planeta Plutón precisa de 250 de nuestros años terrestres para cerrar su órbita en torno al Sol. Y hablando del Sol, su diámetro es de unos 1,5 millones de kilómetros y, su masa, de unos 10^{30} kilogramos. Comparada con él, la Tierra es una pluma de tan sólo 10^{24} kilogramos. No deberíamos olvidar que algunas estrellas –las gigantes rojas– alcanzan diámetro suficiente para albergar en su interior la órbita de Júpiter, aunque claro, tales estrellas son muy tenues, algo así como algodón de azúcar a escala cósmica. En contraste, otras estrellas –las estrellas de neutrones– están tan densa y compactamente aglutinadas que si pudiéramos extraer de ellas un cubo de sólo un milímetro de lado, su masa sería del orden de medio millón de toneladas, igual a la masa del mayor petrolero que se haya construido, a plena carga.

Estos números desbordan tanto nuestra capacidad de comprensión que es virtualmente imposible asombrarnos más. Son números genuinamente fuera del ámbito de lo inteligible, salvo que haya podido lograrse una viva sensibilidad al valor de los exponentes. E incluso dotados de tal sensibilidad resulta difícil reconocerle al universo cuanto le es debido, siendo por una parte tan enormemente grande y, al mismo tiempo, de grano tan sumamente fino.

Pocos minutos después de haber terminado yo un borrador de este artículo recibía, en relación con él, un golpe particularmente duro. Estaba leyendo el periódico cuando me tropecé con una noticia acerca de los gases militares dirigidos contra el sistema nervioso. Decíase que el presidente Reagan esperaba que los gastos a ellos dedicados alcanzasen unos 800 millones de dólares en 1983, y unos 1400 millones en 1984. Me sentí indignado y, al mismo tiempo, casi agradecido de que no fuesen 10.000 millones o 100.000 millones. Sentí de pronto vergüenza de mí mismo. ¿Cómo pude quedar tranquilizado por unos "meros" 1400 millones? ¿Cómo pudieron desgajarse tanto mis pensamientos de la realidad subyacente? Mil millones de dólares para gas nervioso no es solamente lamentable: es odioso. No podemos consentirnos ser más indiferentes todavía de lo que ya somos. Es necesario que nos propongamos salir a empellones de nuestra apatía, porque los "chistes" de esta clase son de un gusto execrable.

En esencia, combatir la estupidez, la memez numérica, no es tan difícil. Requiere fundamentalmente habituarse a la idea de que los números *pequeños*, concretamente, los comprendidos entre 5 y 20 tienen un segundo grupo de significados al ser empleados como exponentes. Sin duda nos parecería revolucionario que los periódicos adoptaran el convenio de expresar números grandes mediante potencias de 10; empero, saber que un número tiene 12 ceros es *más* concreto que saber que se llama billón.

Me pregunto cuántos de nuestros compatriotas se apercibirían al serles dados los números 314.159.265.358.979 y 217.828.182.845 de que el primero es más de 1000 veces mayor que el segundo. Me temo que la mayoría no alcanzarían a apreciarla; me temo también que tampoco serían capaces de leer tales números en voz alta. De ser como temo, tenemos un serio motivo de preocupación.

Hay un libro que se esfuerza en luchar contra esta insensibilidad; un libro lleno de humildad ante algunas de las estupefacientes magnitudes que hemos estado comentando. Se titula *Cosmic View: The Universe in Forty Jumps*. Su autor es un maestro holandés ya fallecido, Kees Boeke. En su libro, Boeke nos lleva a un imaginario viaje en imágenes, cuyos cuadros aumentan de tamaño en progresión geométrica, de razón 1:10. Partiendo de nuestro propio tamaño hay 26 pasos ascendentes y 13 descendentes. Tal vez no sea coincidencia que el libro haya sido escrito por un holandés, pues los holandeses son desde antiguo gentes de mentalidad internacional, viviendo como viven en un país pequeño y vulnerable donde confluyen variedad de culturas y lenguas. Boeke concluye, con lo que me parece ser un estilo característicamente holandés, haciendo votos para que el viaje al que nos lleva sirva de ayuda a las personas, para mejor comprender qué lugar ocupan en el esquema cósmico, y de esta forma contribuya a aproximar más a las gentes de este mundo. Como su conclusión me parece de gran elocuencia, me gustaría terminar tomando una cita de ella:

"Así pues, al pensar a escala cósmica, nos damos cuenta de que el hombre, para llegar a ser realmente humano ha de articular en su ser la máxima humildad con la más considerada y cuidadosa aplicación de las fuerzas cósmicas de que dispone.

"El problema, sin embargo, es que el hombre primitivo tiende al principio a servirse de las fuerzas puestas en su mano, ha de articular en su ser la máxima humildad con la más considerada y cuitoda la creciente familia humana, que ha de vivir confinada en el limitado ámbito de nuestro planeta. Es por tanto cuestión de vida o muerte para la humanidad toda que aprendamos a vivir juntos y velar unos por otros sin considerar nuestro nacimiento o nuestra formación. Ninguna diferencia de nacionalidad, raza, credo o convicción, ninguna diferencia de edad o sexo debe menoscabar nuestro esfuerzo de seres humanos para vivir y trabajar en bien de todos.

"Es por ello necesidad urgente que todos nosotros, tanto niños como adultos, seamos educados en este espíritu y encaminados a esta meta. La humanidad tiene claramente señalado el deber de aprender a vivir junta en mutuo respeto y con el propósito decidido de contribuir a la felicidad de todos sus miembros, sin privilegio para ninguno, y es imperativo que la educación se proyecte sobre este plano.

"En tal educación es elemento importante y necesario alcanzar a desarrollar la visión cósmica; y para ello, para alcanzar una visión tan amplia, tan capaz de abarcarlo todo, pudiera servir de ayuda, aunque sea un poco, la expedición que en estos 'cuarenta saltos a través del universo' hemos realizado. De ser así, ¡confiemos en que muchos la emprendan!"

Taller y laboratorio

Cantando bajo la ducha o el narcisismo por nuestra voz

Jearl Walker

Supongo que el lector se habrá percatado ya de que, en la ducha, la peor de las voces suena magníficamente; al menos, a su propietario. La mía, por ejemplo, de la que en absoluto puede decirse que pertenezca a alguien especialmente dotado para el arte del canto, resulta deliciosa, hasta el punto de que a veces me hago ilusiones. Y es que el recinto de la ducha se muestra muy amable y generoso con los sonidos musicales. Un amigo mío solía practicar allá dentro sus lecciones de violín.

La razón oculta de por qué en el recinto de una ducha mejora tanto la calidad de una ejecución musical hay que buscarla en la resonancia acústica, fenómeno éste que refuerza notablemente la intensidad del sonido. Al igual que los demás sonidos, los generados por los órganos de fonación humanos se propagan por el aire bajo la forma de ondas de presión variable. Si la voz humana pudiera emitir ininterrumpidamente en la misma frecuencia (o tono), dicha onda podría representarse mediante una curva senoidal. En ella se verían gráficamente las variaciones sucesivamente iguales de la presión del aire a lo largo de la trayectoria de la onda sonora, entre un valor relativamente alto (mayor que la presión atmosférica) y otro relativamente bajo (menor que la presión atmosférica).

Tales variaciones son resultado de las pequeñas oscilaciones que ejecutan las moléculas de aire paralelamente a la trayectoria de la onda. En virtud de tales oscilaciones, en algunos lugares la densidad del aire aumenta bastante (y la presión crece), mientras que, en otros, disminuye bastante (y la presión decrece). El conjunto de esa distribución de presiones altas y bajas se propaga desde la boca a la velocidad del sonido, que es del orden de 340 metros por segundo a la temperatura ambiente.

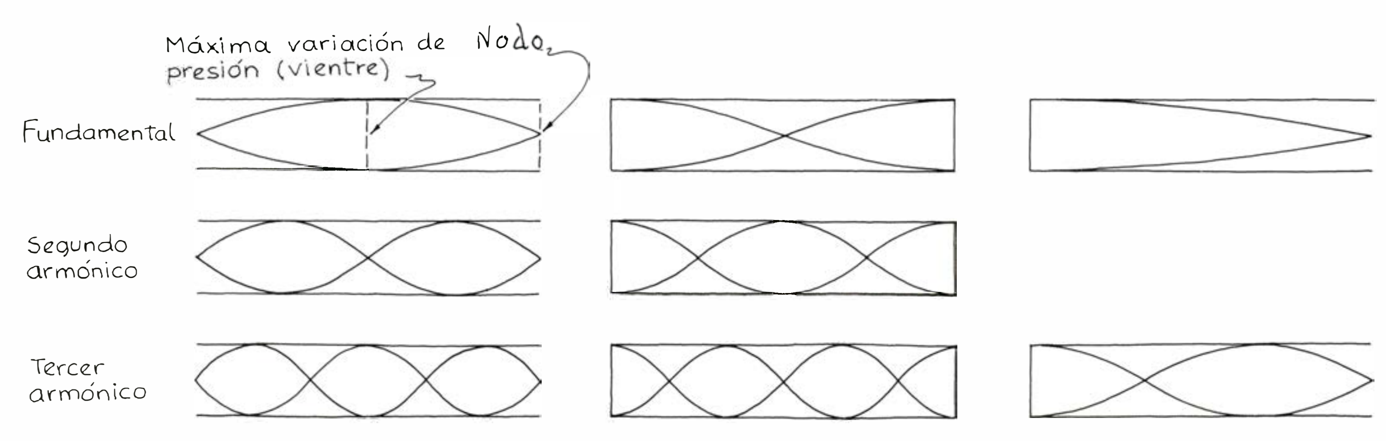
Un parámetro relativo al proceso es la frecuencia del sonido, o número de veces por segundo que pasa un pico de presión de la onda por cada punto de la trayectoria. Así, si ese número es 500, la frecuencia será de 500 hertz (ciclos por segundo). Los límites de sensibilidad auditiva para un adulto joven se sitúan entre los 20 y los 20.000 hertz.

Se llaman ondas progresivas aquellas que viajan desde un punto en dirección a otro. En lo concerniente al canto y resonancia acústica nos interesan, además, otro tipo de ondas: las estacionarias. Consideremos un tubo cilíndrico con ambos extremos abiertos. En primera aproximación, su diámetro no afecta a la resonancia, pero su longitud sí. Supongamos que hacia el interior del tubo se envía sin parar una onda progresiva, e imaginemos que viajamos a lo largo del mismo solidariamente con un pico de la onda. Pese a ser abierto, en el extremo del tubo se refleja hacia atrás parte del sonido, y el resto prosigue hacia el exterior. Entonces la parte reflejada se superpone a las ondas que aún avanzan hacia el final, interfiriendo con ellas. De ese modo, en un instante dado, la presión de un determinado volumen de aire tenderá a aumentar, a causa de la onda que progresa en un sentido, y a disminuir a causa de la que viaja en sentido contrario. Como consecuencia, se genera una distribución de presiones complicada y continuamente mudante. Por otra parte, el sonido que emerja del tubo no será más intenso que el que penetró en él.

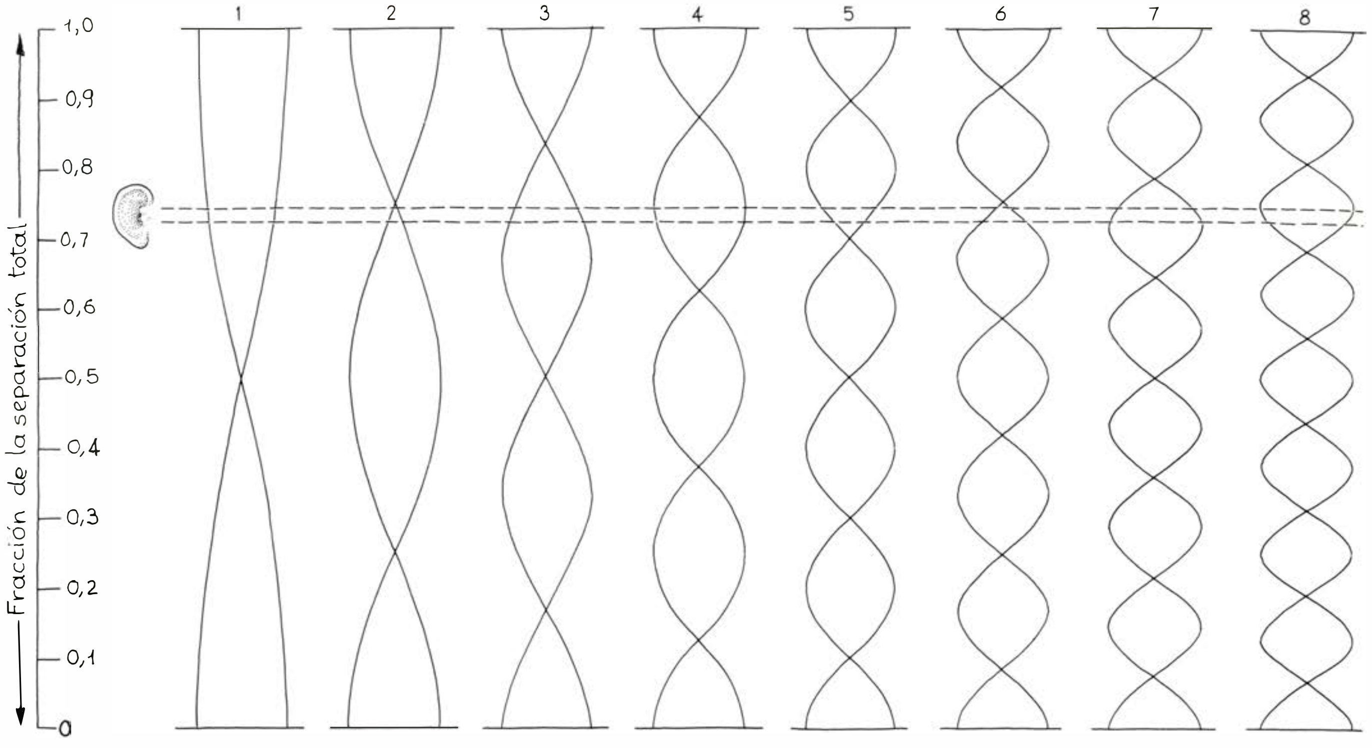
Para determinadas frecuencias se genera una distribución de variaciones de presión mucho más sencilla. Tal distribución se corresponde con las frecuencias propias, o armónicos, del tubo. Estas son todas múltiplos enteros de la menor de ellas, o armónico fundamental. En tal caso, la distribución de variaciones de presión es la de una onda estacionaria, sucediendo que, al crearse una de tales ondas, el sonido que sale del tubo es más intenso que el que penetró en él. Este fenómeno se conoce como resonancia acústica.

Supongamos que se envía al tubo la frecuencia fundamental. Las ondas que avanzan en sentidos contrarios se interfieren; podemos estar, pues, seguros de que en los extremos del tubo (donde se reflejan las ondas) la presión no variará nunca. Esa clase de puntos se llaman nodos. Donde la presión sufre la variación máxima es en el centro del tubo, pasando sin brusquedad de alta a baja, y viceversa (alta cuando ambas ondas interactúan aumentando la presión, baja cuando interactúan disminuyéndola). Los puntos de esta clase reciben el nombre de vientres. La distribución de variaciones de presión presenta una configuración de onda estacionaria, nombre que recibe porque nodos y vientres permanecen inmóviles.

En esta página se ofrece la represen-



Resonancia acústica en tres tipos de tubos



Los ocho primeros armónicos entre el suelo y el techo del compartimento de una ducha

tación gráfica habitual de algunas de las ondas estacionarias posibles en el interior de un tubo. Los puntos donde se cortan las curvas son los nodos y, vientres, los puntos donde la separación entre ellas es máxima. (Esta ilustración puede despistar al sugerir que las moléculas oscilan perpendicularmente al tubo, cuando la realidad es que lo hacen paralelamente a la longitud del mismo.)

Conforme se vayan enviando al tubo frecuencias más altas de la serie armónica, la situación se irá complicando. Con el segundo armónico (frecuencia doble que la fundamental), la onda estacionaria presentará tres nodos, uno en cada extremo y otro en el centro, y dos vientres, uno entre cada par de nodos. El tercer armónico (frecuencia triple que la fundamental) posee cuatro nodos y tres vientres.

La importancia de la resonancia acústica radica en que, con ella, se crea una gran onda estacionaria, más intensa que el sonido enviado al tubo, de modo que éste actúa de amplificador. En la explicación elemental que acabo de ofrecerles, los valores de la serie armónica de un tubo dependen únicamente de la longitud de éste y de la velocidad del sonido. Caldeando el aire, con lo que aumenta la velocidad del sonido, los armónicos se corren hacia arriba. Acortando el tubo se produce el mismo efecto, razón por la cual el armónico fundamental de los tubos de órgano cortos es más alto que el de los largos.

Hasta aquí sólo hemos hablado de tubos de extremos abiertos, pero en lo que respecta al canto y la ducha hay que considerar otros dos dispositivos resonadores. Uno de ellos es el tubo de extremos cerrados, cuya serie armónica es igual a la del tubo de extremos abiertos, aunque sean diferentes las ondas estacionarias. En efecto, en este caso, en los extremos se forman vientres porque las paredes imponen que, en ellas, deben darse las variaciones de presión máximas. La representación gráfica del primer armónico (o fundamental) muestra un vientre en cada extremo y un nodo en el centro.

Cuando el tubo tiene un sólo extremo cerrado, la serie armónica se modifica sutilmente. Un dibujo del primer armónico sigue ateniéndose a las reglas que les he resumido, exhibiendo un vientre en el extremo cerrado y un nodo en el abierto. Lo mismo acontece con el armónico siguiente; mas aquí la frecuencia es triple que la del primer armónico, o fundamental. A veces, esta onda estacionaria recibe el nombre de segundo armónico, porque es la segunda de las frecuencias más bajas posibles en un tubo como el considerado. Esa denominación confunde más que aclara, y así seguiré hablando de tercer armónico, dado que su frecuencia triplica la fundamental. En la resonancia dentro de los tubos de un extremo abierto sólo pueden presentarse los armónicos impares (primero, tercero, quinto, etc.). Las reglas que gobiernan las variaciones de presión en los extremos de los tubos no admiten armónicos pares.

En el arte del canto hay que conseguir un efecto de resonancia dentro de lo que es esencialmente un tubo de un extremo abierto. No se trata del cuerpo recto y cilíndrico que hemos venido considerando, desde luego; pero el modelo sirve perfectamente. Se genera el sonido cuando el aire espirado por los pulmones atraviesa las cuerdas vocales (que no son tales cuerdas, sino membranas delgadas). El aire emerge de las cuerdas vocales en una sucesión de pulsos, cuya frecuencia depende fundamentalmente de la tensión de aquéllas; a mayor tensión, mayor frecuencia de las descargas de aire y, por tanto, mayor frecuencia.

Las cuerdas vocales vibran según una serie armónica de frecuencias, generando correspondientemente una serie armónica de ondas sonoras, que comprende el armónico fundamental y todos los superiores. De estas ondas, la más intensa es la fundamental.

Dichas ondas atraviesan la cavidad, o tracto vocal compuesto de la laringe, la faringe y la boca. El tracto vocal es asimilable a un tubo de un extremo cerrado (por las cuerdas vocales) y otro abierto (la boca); por tanto, le corresponderá una serie armónica a cuyas fre-

cuencias podrá incitarse la aparición de ondas estacionarias. Entre las frecuencias que penetren en el tracto vocal procedentes de las cuerdas se encontrarán algunas iguales o próximas a aquellas frecuencias de resonancia. Estas frecuencias son relativamente intensas.

Un cantante generará un sonido de la máxima intensidad cuando sean iguales los armónicos de las cuerdas vocales y los del tracto. Esta igualación, o adaptación del tracto vocal a los armónicos de las cuerdas vocales, la consiguen los buenos cantantes de varias maneras, posiblemente con un esfuerzo consciente mínimo. La tensión de las cuerdas vocales puede corregirse un tanto para ajustar el armónico fundamental en dicho punto; pero, una vez establecido dicho armónico, la adaptación final debe conseguirse alterando la forma del tracto vocal.

En este caso la física de la resonancia acústica no es tan simple como en los tubos ideales, para los que he pasado por alto el efecto del diámetro. Haciendo variar convenientemente la forma de su tracto vocal, todo varón adulto puede conseguir, al cantar, que el armónico fundamental del tracto adquiera cualquier valor comprendido entre 250 y 700 hertz. El tercer armónico (téngase presente que, en este tipo de tubo, sólo se generan armónicos impares) varía entre 700 y 2500 hertz.

Mi aptitud para adaptar el tracto vocal al armónico fundamental que generan mis cuerdas vocales es nula. Así, si me propusiera lanzar a un auditorio una nota de, pongamos, 500 hertz tendría que gritar para hacerla audible. Sin embargo, un cantante adecuadamente preparado conseguiría la adaptación y aprovecharía la amplificación debida a la resonancia en el tracto vocal. La ventaja de cantar en la ducha reside en que al cantor inexperto le ayuda la resonancia generada entre las superficies del recinto.

El efecto del compartimento de la ducha es el de un tubo con los extremos cerrados. No obstante, se diferencia de los tubos que les he descrito en que posee tres pares de extremos: (1) el suelo y el techo, (2) una pareja de paredes paralelas y (3) la otra pareja de paredes paralelas. (Supondré que la cortina o la puerta son una pared maciza.) Entre dos paredes cualesquiera puede provocarse la aparición de ondas estacionarias. Una representación gráfica de la onda fundamental correspondiente, por ejemplo, a la pareja suelo-techo muestra vientres en el suelo y en el techo y un nodo a medio camino entre ambos. La del segundo armónico (en los tubos con los extremos cerrados puede haber armónicos pares) posee tres vientres (en el suelo, techo y a medio camino) y dos nodos (entre los vientres). Para las otras parejas de paredes se dan configuraciones similares.

Aunque la relación entre el canto en la ducha y la resonancia acústica se conoce desde hace tiempo, el único trabajo acerca del tema que he podido localizar es el resumen de una charla ofrecida por Daniel W. Haines, de la Universidad de Carolina del Sur, en la Segunda Conferencia sobre Enseñanza de Acústica y Física del Sonido en abril de 1976. Este trabajo revela algunos aspectos muy interesantes del fenómeno. Por ejemplo, supóngase que nos hayamos propuesto generar resonancia entre el suelo y el techo del compartimento de la ducha. Si la fuente sonora se halla a medio camino entre ambos, no habrá resonancia porque ese punto corresponde a un nodo de la onda estacionaria asociada. En efecto, no podrá darse un nodo, que por definición es un punto donde la variación de presión es nula, allá donde la propia fuente crea variaciones de presión. Para conseguir lo más parecido a un vientre de una onda estacionaria lo mejor sería cantar junto al suelo o al techo (aunque ello reviste el peligro de ponernos en ridículo ante algún observador casual, por muy de confianza que sea).

Si la fuente sonora se encuentra a medio camino entre el suelo y el techo pueden darse el segundo armónico y otros armónicos pares. Para otras posiciones de la fuente pueden aparecer numerosos armónicos, unos más intensos que otros, según la cercanía de aquél a algún vientre. Los únicos armónicos imposibles son aquellos para los que coinciden la fuente y un nodo. (Mi explicación a este respecto puede ser imperfecta: si el manantial sonoro es muy pequeño e irradia sonido a todos los puntos del recinto, posiblemente su ubicación carezca de importancia. Creo, sin embargo, que la cabeza humana es lo bastante grande para que su posición importe mucho.)

Haines hizo notar asimismo que al cantante puede sonarle bien o mal su propia voz de acuerdo con la posición de sus oídos. Si éstos se encuentran en un nodo, el cantor no percibirá la onda estacionaria (ya que la audición la provocan las ondas de presión al actuar sobre los tímpanos). Si se encuentran en un vientre, la impresión sobre los tímpanos será máxima.

En la ilustración de la página precedente se representan los ocho primeros armónicos entre el suelo y el techo. La boca y los oídos del cantante se hallan aproximadamente a un cuarto de la distancia entre suelo y techo. Teóricamente le es posible desencadenar y también oír el armónico fundamental, al no tener la boca ni los oídos en un nodo de la onda estacionaria correspondiente. Como tampoco están en un vientre, no podrá desencadenar o percibir intensamente un armónico fundamental. Al tener la cabeza en uno de los nodos del segundo armónico, no podrá desencadenar éste ni percibirlo. De las ondas estacionarias representadas, las que mejor percibiría serían las del tercero, cuarto, séptimo y octavo armónicos.

Respecto a las ondas estacionarias generadas entre dos paredes, el cantante goza de mayor facilidad de movimiento. Alejándose de una pared, o acercándose a ella, puede modificar el espectro de armónicos que genera o

	De suelo a techo: 2,095 metros	De pared a pared: 1,375 metros	De pared a pared: 1,010 metros
Armónico	Frecuencia (hertz)	Frecuencia (hertz)	Frecuencia (hertz)
1	83	126	171
2	165	252	343
3	248	377	514
4	330	503	685
5	413	629	856
6	495	755	1028
7	578	881	
8	661	1007	
9	743		
10	826		
11	908		
12	991		
13	1074		

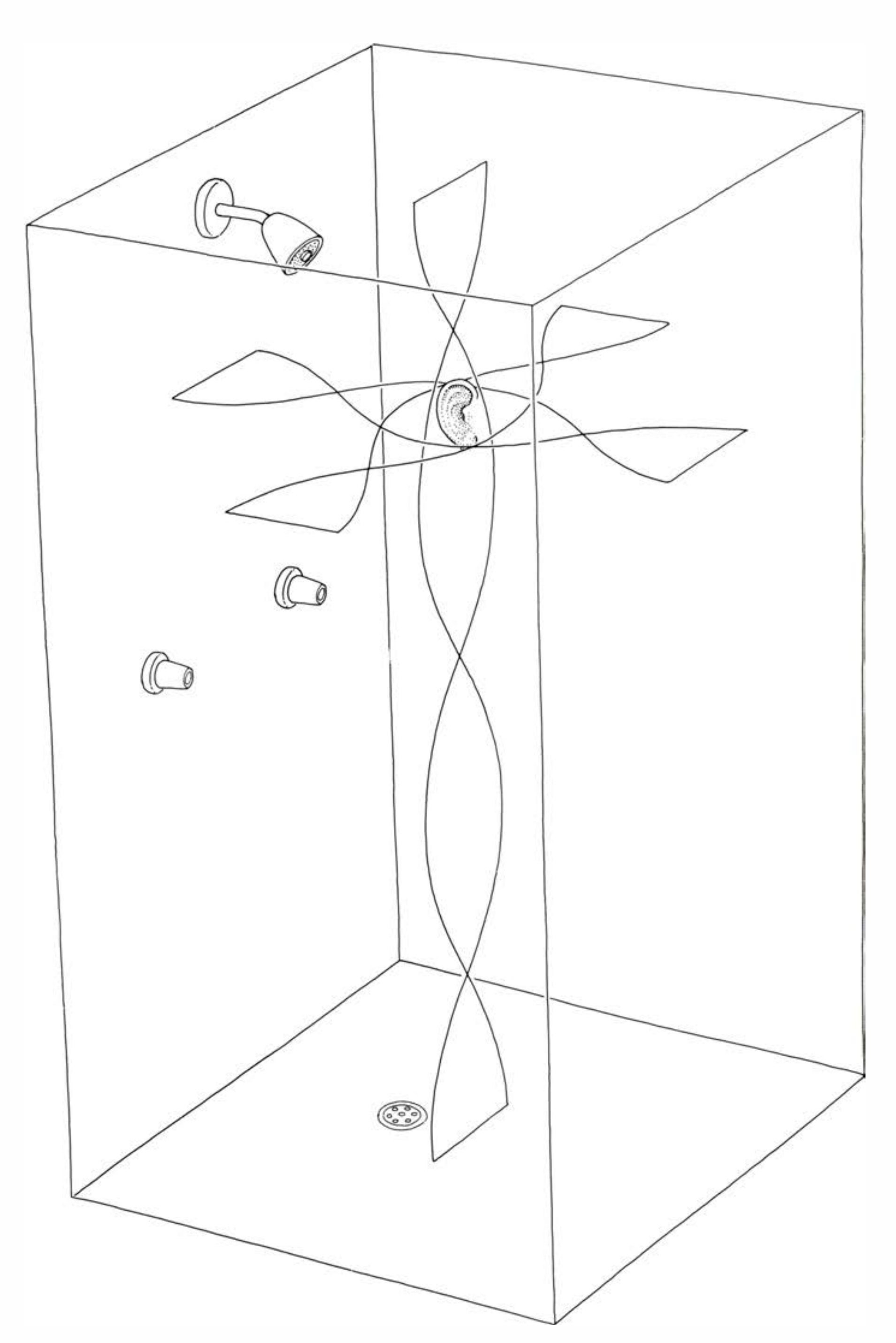
Armónicos en el compartimento de la ducha

percibe. En el centro del compartimento estará ausente el armónico fundamental de los dos pares de paredes, pues a ambos corresponde un nodo en ese punto; sin embargo, el segundo armónico posee un vientre allí y entrará en juego. Acercándose a una de las paredes, el cantante eliminaría el segundo armónico, pero desencadenaría el tercero.

La separación horizontal entre boca y oídos complica los armónicos entre las paredes. Para determinados armónicos, puede que la boca se encuentre en la posición adecuada para generar una onda estacionaria, mientras que los oídos estén en un nodo. (Para estudiar el caso correctamente debe también considerarse la distorsión que la cabeza produce sobre la onda estacionaria.)

Surge una complicación más a causa de que las frecuencias de los armónicos no son exactamente como se indica en la ilustración de la página precedente. Los valores del cuadro dependen de la velocidad del sonido, la cual puede que varíe conforme la ducha vaya calentando el aire. E incluso sin considerar este hecho: las frecuencias no tienen valores definidos tajantemente, sino que pueden oscilar entre márgenes de unos 10 hertz. El modo más eficaz de generar una onda estacionaria es que el cantante entone la voz a una de las frecuencias reseñadas en el cuadro; si se desvía levemente de ese valor teórico, la onda será más débil.

Otra complicación resulta del hecho de que el cuerpo del cantante ocupa un espacio dentro del compartimento y refleja parte del sonido. Lo mejor será suponer que ello es secundario respecto al mecanismo de creación de ondas estacionarias. Estudiando las resonancias como si el compartimento de la ducha fuese un tubo de extremos cerrados se consigue por los menos estimar aproximadamente los valores de las frecuencias de resonancia.

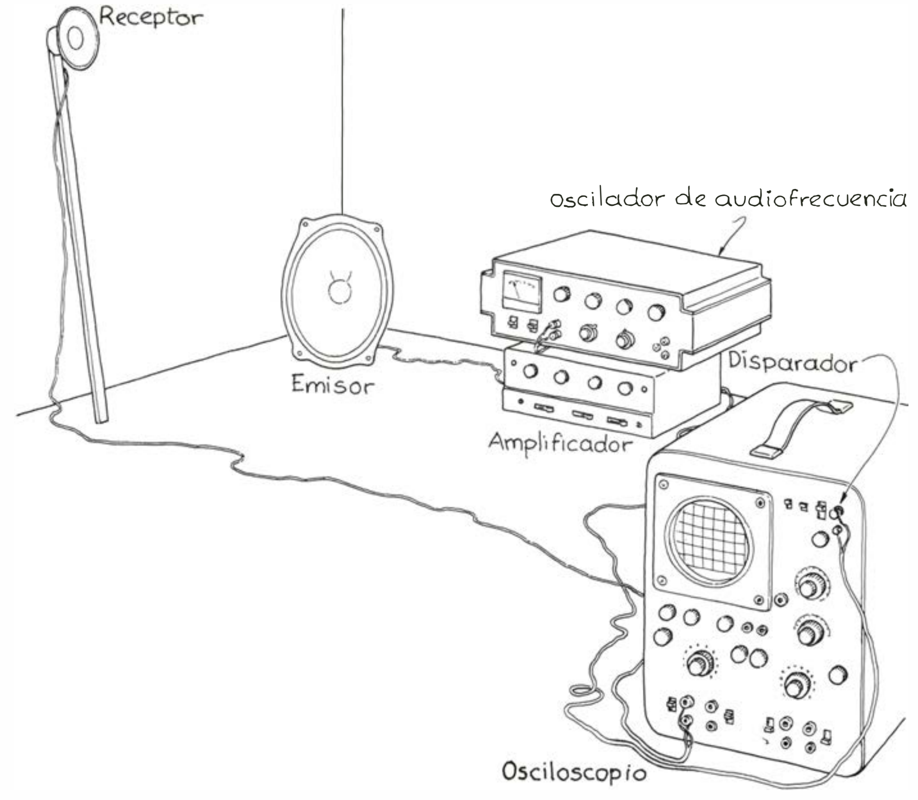


Configuración de algunos armónicos en el recinto de una ducha

Las frecuencias que aparecen en el cuadro de la ilustración de la página precedente las he recogido en mi ducha, cuyo compartimento tiene paredes, suelo y techo recubiertos de azulejos; una de las paredes es una puerta de vidrio. Las he calculado bajo la hipótesis de que, en aquel recinto, la velocidad del sonido es de 346 metros por segundo. Así, cada frecuencia fundamental será esa velocidad dividida por el doble de la distancia entre cada pareja de superficies reflectoras.

Como el suelo y el techo están mucho más separados entre sí que las paredes, la frecuencia fundamental correspondiente a ellos será menor que la correspondiente a las paredes. Los armónicos superiores se calculan multiplicando la frecuencia fundamental por números enteros; por consiguiente, el segundo armónico es doble que la frecuencia fundamental, triple el tercero, y así sucesivamente.

Supongamos ahora que me pongo a cantar en la ducha, ajustando mis cuerdas vocales a una frecuencia fundamental de 330 hertz, pero que no consigo adaptar a ésta mi tracto vocal. Mis cuerdas vocales generan frecuencias que son múltiplos enteros de la fundamental, o sea, 330, 660, 990 hertz, etcétera. Supongamos que mi tracto vocal esté ajustado a 450 hertz, cuya frecuencia de resonancia siguiente (tercer armónico) es de 1350 hertz. Fuera de la ducha mi canto sonará débil, puesto que ninguna de las frecuencias procedentes de las cuerdas vocales resonarán en el tracto. Dentro de la ducha, por el contrario, las frecuencias de 330 y 660 hertz crearán ondas estacionarias entre el suelo y el techo, excitando la menor de ellas el cuarto armónico y la mayor el octavo armónico. (El sonido de 990 hertz desencadenaría el doceavo armónico, pero mi boca se encontraría en un



Montaje electrónico para captar armónicos

nodo de la onda correspondiente.) Por tanto, me oiré maravillosamente, pese a la falta de resonancia en el tracto vocal.

Cuando me dispuse a experimentar en torno a los armónicos que se desarrollaban en el recinto de mi ducha me encontré con mi incapacidad para sostener una nota largo tiempo y con mi falta de pericia para determinar su frecuencia; por eso, en lugar de mi voz utilicé un oscilador de audiofrecuencia y un altavoz. La frecuencia de la señal podía gobernarla actuando sobre el mando del oscilador. Esa señal, tras aumentarla en un amplificador, la dirigía a un cono de altavoz cuya ubicación dentro del compartimento podía variar a voluntad. La apertura del cono era de 19 por 12 centímetros; uno más pequeño no habría emitido las frecuencias bajas.

A quien se decida a efectuar experiencias de este tipo debo recordarle que los artefactos eléctricos son peligrosos en presencia de agua. Otro riesgo es que la intensidad de alguna de las ondas estacionarias desencadenadas en el compartimento pueda resultar excesiva y lastimar el oído, especialmente cuando se eleve el volumen del altavoz. Por cuya razón empleé unos auriculares amortiguadores de sonido.

Sintonizando el oscilador a lo largo de las frecuencias comprendidas entre los 200 y los 1000 hertz me fue posible percibir la aparición y desaparición de las resonancias. Cuando el oscilador no resonaba con el recinto, el sonido dentro de éste se sentía relativamente débil; pero, en las resonancias, su intensidad crecía, a veces espectacularmente.

Fijando el oscilador en alguna de las frecuencias de resonancia, podía hacer variar la intensidad con que percibía el sonido trasladando la cabeza por todo el recinto. Cuando mis oídos pasaban desde un vientre a un nodo, el nivel de sonido decaía abruptamente. Con dos o tres movimientos de tanteo pude determinar si la onda estacionaria se encontraba entre dos paredes (y, en tal caso, entre qué dos paredes) o entre el suelo y el techo.

La efectividad con que el altavoz generaba una onda estacionaria dependía en parte de la orientación que le daba. Así, al orientarlo hacia arriba únicamente conseguía una onda estacionaria débil entre las paredes, aunque estuviera situado correctamente. Descubrí también que si lo ponía inclinado en una esquina sobre el suelo, el sonido que difundía podía considerarse orientado, más o menos correctamente, para generar todas las ondas estacionarias posibles.

Para calcular de una manera más objetiva las frecuencias de resonancia instalé un osciloscopio y un segundo altavoz; éste algo más pequeño que el altavoz emisor. (Un altavoz chico no habría conseguido captar las frecuencias más bajas.) El altavoz interceptaba parte de las ondas sonoras del recinto y creaba la representación gráfica correspondiente en la pantalla del osciloscopio. La componente vertical de esa figura representaría así la amplitud de la onda sonora; en el eje horizontal estaba el tiempo de barrido del osciloscopio.

Conecté los bornes de entrada del circuito disparador del osciloscopio al oscilador de audio, de modo que el barrido de un lado a otro de la pantalla estuviera sincronizado con la salida de éste. Al establecer la frecuencia del oscilador en coincidencia con un armónico del compartimento de la ducha, la figura dada por el osciloscopio aumentaba de amplitud. Medí la gama de frecuencias generadoras de resonancia determinando las dos frecuencias (una superior a la óptima y la otra inferior a ésta) para las que la amplitud de la figura se reducía a la mitad de la máxima.

Para trazar la gráfica de las ondas estacionarias del compartimento desplacé el altavoz receptor entre las paredes tanto horizontal como verticalmente. Para localizar un armónico entre suelo y techo, apuntaba el altavoz hacia arriba y lo movía verticalmente. Para localizar ondas estacionarias entre dos paredes, lo apuntaba hacia una de ellas y lo movía adelante y atrás horizontalmente. Así podía asegurarme de si una resonancia correspondía a una onda estacionaria entre las paredes o entre el suelo y el techo. Para reducir a un mínimo la distorsión que pudiera introducir mi cuerpo, afirmé con cinta adhesiva el altavoz a una vara de medir, que manejaba por el interior del compartimento; pero, aun así, mi cuerpo y los aparatos distorsionaban las ondas estacionarias lo bastante para que pudiera notarlo.

Salvo algunas de las frecuencias más bajas, a las que eran insensibles los altavoces, pude hallar e identificar la mayoría de los armónicos de mi cuadro de valores. La frecuencia a la que conseguía la mejor curva en el osciloscopio no siempre coincidía con el valor teórico, pero su registro solía abarcar los valores del cuadro. Me hubiera agradado observar el corrimiento de los armónicos con el aumento de la temperatura del aire. Mas, para ello, hubiera tenido que abrir el grifo del agua caliente, lo que hubiera supuesto un peligro de electrocución excesivo.

Las personas percibimos los armónicos en el recinto de una ducha de mane-

ra diferente a como lo estimé yo, más objetivamente, con el concurso de artefactos electrónicos. Efectivamente, el oído acusa las frecuencias puras como si estuvieran formadas por otras varias. Ello se debe a que el sistema auditivo no reacciona puntualmente a las ondas sonoras. Por eso se dice que el sistema auditivo es alineal, queriendo indicar que, al recibir y procesar una señal, la deforma.

La alinealidad del proceso auditivo posibilita que haya personas que perciban frecuencias "irreales". Si suponemos que al oído llegan dos frecuencias el oyente no percibirá esas dos frecuencias sólo, sino que tal vez aprecie también la diferencia entre ambas. Por ejemplo, si las dos señales son de 500 y 650 hertz, quizá se perciba una señal de 150 hertz. Sean ahora f_1 y f_2 las dos frecuencias realmente emitidas; el oyente puede asimismo percibir tonos de $2f_1-f_2$, $3f_1-2f_2$, etcétera. O sea, pueden apreciarse señales de baja frecuencia cuando, en realidad, al oído sólo llegan frecuencias superiores.

Cabe que tales frecuencias, llamadas tonos de combinación, enriquezcan nuestro canto mañanero. Supongamos alguien que cante con el armónico fundamental de sus cuerdas vocales a 289 hertz. Los múltiplos enteros de éste serán 578, 867, etcétera. Como siempre, nuestro cantor tampoco consigue adaptar su cavidad vocal a aquel armónico. Supongamos que, en su lugar, el tracto ofrezca un armónico fundamental de 330 hertz y, por ello, un tercer armónico de 990 hertz.

El sonido de 289 hertz lo amplificará el tracto vocal muy deficientemente, y el de 589 hertz aún peor, con lo que el canto sonará débil. El sonido de 578 hertz, sin embargo, excita el séptimo armónico entre suelo y techo. Entonces, puede que así haya algún tono de combinación que refuerce el sonido de 289 hertz. Efectivamente, la diferencia entre la frecuencia de la señal de 289 hertz (débilmente amplificada en la cavidad vocal) y la de 578 hertz (amplificada por el recinto) es 289 hertz y, si el cantor es capaz de apreciar un tono de combinación, la señal de 289 hertz le sonará más intensa de lo que es en realidad.

Dejaré para el lector ulteriores investigaciones acerca de los tonos de combinación. Si alguien encuentra efectos de interés, será un placer para mí recibir sus noticias. Me agradaría, asimismo, recibir aviso de posibles fallos en mi línea de razonamiento acerca de las causas de que, casi siempre, todos cantemos tan bien en la ducha.

Libros

Sobre ecología teórica y aplicada, progreso científico y antropología

Antonio Escarré, Oriol de Bolós, Mariano Artigas y Claudi Esteva

ECOLOGÍA, por Ramón Margalef. Editorial Planeta, Barcelona, 1981; 252 páginas. Este libro merece una anotación preliminar: no se trata de ninguna novedad en el sentido editorial riguroso. Aparece como por una especie de proceso de gemación a partir de la Enciclopedia Temática Planeta, en unos momentos en que, ya que casi todo se puede hacer "más ecológicamente", la empresa ha debido prever un mercado inusual en un libro de ciencia con este binomio título-autor, mucho más desvaído comercialmente como bloque de una enciclopedia. A esta obra se le conoce una hermana, escrita por la misma época y editada por la Universidad Nacional de Educación a Distancia.

Los ecólogos son científicos casi necesariamente abocados al generalismo. De esta tendencia ha comentado Margalef con cierto humor que, mientras que los especialistas corren el riesgo de saber muchísimo de prácticamente nada, los ecólogos se verían en peligro de saber casi nada de prácticamente todo. En el autor de esta obra no se cumple el fatal destino de los ecólogos, pese a su reconocida tendencia hacia una ecología de síntesis que busca poner de relieve las regularidades que rigen el funcionamiento de la naturaleza.

La lectura del libro me ha llevado casi inconscientemente a la voluminosa "Ecología" de Ramón Margalef (Editorial Omega, 1974, 951 páginas), en parte por ese vicio de estudio comparativo de estructuras propio de los naturalistas y, en parte, por un párrafo de su prólogo en el que declaraba haber hecho un libro "motivado y parcial"... porque... "la Universidad necesita creatividad y cierta dosis de provocación a la polémica".

Observamos una estrecha correspondencia entre los contenidos de ambas obras, que no es en absoluto achacable a que se siga un determinado *corpus* temático, el más seguro para los textos de ecología. El apartado inicial, intitulado "La ecología, la tierra y la vida", viene a ser un extracto de las partes primera y segunda (el Medio y la Biogeografía en la obra mayor), con un apéndice sobre demografía, pero analizada desde el punto de vista de la evolución. Bajo el epígrafe "El funcionamiento de los ecosistemas" se recogen aquí toda la ecología trófica, el resto de la ecología demográfica, es decir, el sistema depredador-presa (divulgado como modelo de Tom y Jerry) y la competencia, así como un breve apartado sobre diversidad, al que queda reducida la ecología descriptiva. El tercer y último apartado, sobre "La biosfera y el hombre", es el más extenso, y se corresponde con los capítulos dedicados a la sucesión y explotación humana, regresión y conservación, que leemos en el tratado de Omega.

A la hora de hacer un balance de la selección llevada a cabo en la confección de esta obra menor cabe aplicar la declaración de intenciones a que antes me referí. Se recopilaba allí una información bastante exhaustiva; las interpretaciones sobre los hechos expuestos quedaban un poco perdidas entre una exposición de datos que resultaba algo abrumadora a veces. Aquí, por contra, tenemos un resumen mucho más dominado por la visión personal que Margalef tiene de la ecología.

Este sesgo subjetivo del libro hace que su lectura resulte un apasionante bombardeo de ideas originales sobre las reglas que rigen el funcionamiento de la biosfera. Creo reconocer un esfuerzo divulgador que se traduce en un estilo más sencillo y en la reducción al mínimo, tanto de expresiones matemáticas como del uso de la nomenclatura sistemática, que tan poco digerible resulta para los no iniciados, debido a la enorme diversificación de las formas de vida que, según el propio autor, cabe calificar tanto de admirable como de desesperante. No obstante el esfuerzo vulgarizador, el libro resulta denso y pide una lectura reflexiva.

Hay varios aspectos que merecen subrayarse. En primer lugar, cuando describe las cubiertas fluidas del planeta incluye un esquema especialmente afortunado desde el punto de vista pedagógico, donde se asimila el conjunto formado por la atmósfera y la hidrosfera a una máquina térmica. También es convincente el tratamiento que se da a la interacción radiación-materia viva, con la selección de las longitudes de onda acordes con las distancias de enlace en las moléculas orgánicas, y la explicación del límite superior de la producción primaria con una analogía por medio de un modelo hidráulico.

Otro interesante problema de límite es el de la diversidad de especies en una comunidad. El autor propuso, hace unos veinticinco años, medir la diversidad de una biocenosis de la misma forma como se evalúa el contenido en información de un sistema. Sugerencia que siguieron un buen número de ecólogos, lo que ha permitido reunir en este tiempo una cantidad muy considerable de estimas sobre taxocenosis de ecosistemas muy diversos. Con ello se ha puesto de relieve la existencia de un tope máximo de diversidad, del orden de los 5 bits/individuo, que se interpreta como una consecuencia de la organización jerárquica y funcional de los ecosistemas. Resulta atrayente, por último, la presentación de los distintos tipos de ecosistemas (mar, ríos, lagos, ecosistemas terrestres), así como la inclusión de la acción del hombre sobre los mismos, siempre con el fenómeno de la sucesión ecológica como telón de fondo, tema en el que el autor ha hecho importantes aportaciones para el desarrollo de la ecología. (A. E.)

ECOLOGÍA Y PAISAJE, por F. González Bernáldez. H. Blume Ediciones. Madrid, 1981; 255 págs. En los tiempos presentes geógrafos y ecólogos se interesan vivamente por el concepto de paisaje, que hasta hace pocos años era de uso casi exclusivo de los artistas. La rápida deterioración del entorno, de la que son a la vez autores y testigos las generaciones actuales, ha contribuido de modo especial a aumentar la preocupación por las complejas realidades del ambiente y por el modo como el hombre resulta afectado por ellas.

El libro que comentamos, denso y rico en ideas, se extiende en el examen de lo que es el paisaje y de las relaciones entre el mismo y la vida humana. Su lectura ha de ser útil, especialmente para inducir a reflexión a aquellas personas que deben ocuparse del paisaje en los distintos aspectos, lo mismo hombres de ciencia o técnicos forestales y agrónomos que urbanistas o educadores. El texto intenta incluir, dentro de una visión racional y científica, también los aspectos estéticos y emociona-

les que, subyacentes siempre en la valoración de los paisajes, son a menudo menospreciados o tratados de modo subjetivo y arbitrario.

Ante todo, el autor discute las definiciones de paisaje. Bajo este nombre se incluyen principalmente dos acepciones fundamentales: por un lado se considera la imagen del territorio con sus diversidades y por el otro se quiere expresar el sistema de elementos ligados por relaciones de interdependencia que existe en el territorio. Tanto el artista como el hombre de ciencia pueden hablar de paisaje, aunque vean en él aspectos distintos. Existe una cierta tensión, evidentemente, entre los aspectos científico-racionales y los estético-sensoriales del paisaje; la resultante depende de la cultura y civilización en que nos hallemos inmersos. En último término el concepto original de paisaje tiene connotaciones que, lo mismo para unos que para otros, hacen sobresalir su valor como fuente de información. El hombre y sus predecesores llevan muchos miles de años inmersos en el flujo de información que es el paisaje; no resulta fácil, pues, separarse de él. En definitiva, hay que reconocer que el paisaje es un valioso recurso natural, cuyo aprecio requiere a la vez conocimientos (ciencia) y sensibilidad.

En capítulos sucesivos se estudia la estructura del paisaje, formado por elementos que constituyen un sistema con significado global propio, no reducible a la suma del de sus componentes aislados; se expone cómo tales elementos se van repitiendo en forma más o menos reiterativa de modo semejante, según pautas propias de cada caso, y se examina el funcionamiento del sistema resultante. Se trata luego del modo como se puede profundizar en el conocimiento de los paisajes y de su significación. En un capítulo especial se estudia el uso de indicadores, es decir, de señales que permiten conocer el estado del sistema y su evolución. Pueden servir de indicadores ciertos componentes geológicos o biológicos y determinadas trazas de la actividad del hombre. Una atención especial se dedica también a la percepción a distancia (teledetección), que en los últimos tiempos va adquiriendo importancia fundamental gracias a la fotografía aérea y a la utilización científica de satélites. Se discuten los distintos métodos de captación de datos y de interpretación de los resultados. Se detallan los métodos matemáticos de desciframiento de la información y de búsqueda de estructuras.

De la interpretación de los datos y de la ordenación de los mismos resulta la sistemática de las pautas del paisaje. El reconocimiento de tales pautas repetitivas y de las relaciones que existen entre los distintos elementos es fundamental para el estudio integrado de los territorios. Cada porción de la superficie terrestre, con su relieve, su clima, sus aguas, su vegetación, etcétera, puede considerarse un todo natural, formado por componentes interdependientes e interactivos. ¿Cómo puede enfocarse en la práctica el estudio integrado? Es fundamental para ello el reconocimiento de las relaciones que existen entre los componentes y de las pautas de reiteración. En general, suele ser posible la identificación de una jerarquía de pautas repetitivas o *patterns* de distinta categoría, desde detalles muy finos hasta territorios inmensos.

"Paisaje e historia" es el título del capítulo octavo. En general, los paisajes están sometidos a procesos de variación en el tiempo, debidos a cambios naturales o, en los tiempos modernos, a la acción del hombre. Las transformaciones pueden producirse con ritmo muy variado, de modo que los paisajes actuales suelen comprender retazos de época distinta. Para interpretar un paisaje hay que recurrir, pues, a la historia geológica y la historia humana, antigua y moderna. La explicación se ilustra con ejemplos variados. En lo que respecta a la región mediterránea, el autor sostiene que la supresión radical de la intervención del hombre en paisajes ya muy alterados puede en algunos casos provocar degradación, lo mismo que la provoca la explotación abusiva. Sobre este punto nos parece necesario matizar un poco lo que dice el texto, pues en los ejemplos que en él se citan no ha habido tiempo ni posibilidad alguna de que se alcancen estados naturales de equilibrio estable. La cesación de la acción humana lleva ciertamente algunas veces, en las primeras fases de recuperación de la vegetación, a situaciones provisionalmente inestables o incluso peligrosas y no puede excluirse que en determinadas circunstancias conduzca, por ejemplo, al mantenimiento por largo tiempo de comunidades vegetales pirofíticas o incapaces de proteger de modo suficiente el suelo contra la erosión. Pero no hay duda de que, aun en la región mediterránea, si la vegetación logra alcanzar estados próximos al de clímax (encinares, etcétera), su estabilidad natural y la global del ecosistema suelen quedar plenamente aseguradas. En este aspecto un conocimiento adecuado de las consecuencias que puede tener la acción del hombre en los distintos estados de los sistemas ecológicos resulta de urgente y primordial necesidad. A ello se refiere el capítulo noveno, en el que se indica sumariamente también el proceso histórico de humanización del paisaje. En Castilla y Extremadura se distinguen, de acuerdo con el nivel de transformación, la "mancha" o bosque denso, la "dehesa", bosque aclarado empleado para pasto, y, en último lugar, el campo, destinado al cultivo agrícola. Cada uno de los ambientes caracterizados por la vegetación presenta también su fauna particular, a cuyo conocimiento es necesario prestar la mayor atención.

La última parte del libro, una de las más originales, se dedica al estudio de los valores humanos y sociales del paisaje. Se aborda el valor estético y emocional del paisaje. Parece necesario estudiar con mayor atención y con método científico lo que se refiere a los valores "intangibles" de tipo estético, sentimental, etc. Tales valores se hallan siempre subyacentes en las interpretaciones que se dan del paisaje. Ni siquiera los hombres de ciencia que se han propuesto tener en cuenta ante todo el aspecto racional del estudio del paisaje pueden evitar, por lo común, la influencia del aspecto emocional, como lo muestran los ejemplos de Humboldt, de Darwin, etcétera, en cuya obra escrita una parte muy considerable del texto se dedica a la consideración del paisaje en forma extracientífica.

Pero el estudio científico de las calidades estéticas del paisaje está en su infancia. Según el autor, parece que las preferencias estéticas se apoyan en reacciones instintivas de carácter simbólico provocadas por determinados elementos del entorno, reacciones que han llegado a generarse en el curso de millares de años, en virtud de un proceso de adaptación al ambiente. A lo largo de la historia de la humanidad el paisaje ha sido siempre fuente de información útil y debe seguir siendo útil. A ello puede llegarse de forma racional y sistemática mediante la búsqueda de correspondencias entre signos visibles y características menos patentes de funcionamiento de la naturaleza. Pero la apreciación de la "belleza del paisaje" mediante criterios objetivos es realmente difícil. Se ha intentado la evaluación del valor estético de los paisajes, que debe tenerse en cuenta, por ejemplo, al realizar trabajos de ordenación territorial, por medio de expertos, análogos a los críticos de arte, y también mediante consulta popular (encuestas, etcétera). A la teoría formalista, según la cual el valor estético depende de que exista cierta armonía en las combinaciones de formas y de colores, se oponen, probablemente con razón, aquellas teorías que reconocen el papel pre-

dominante de los contenidos expresivos, lo que hace que un mismo paisaje pueda tener valores muy diferentes para distintos espectadores.

La capacidad de aumentar el disfrute estético y emocional del paisaje es enorme. Para ello el conocimiento del sentido de las cosas y la experiencia, al elevar la capacidad de interpretación del entorno, alcanzan un papel de primer orden. La finalidad de la educación ambiental es facilitar la comprensión de la compleja estructura del medio ambiente y la adquisición de conocimientos que permitan resolver los problemas que presenta la naturaleza inmediata. En el libro se exponen las modalidades de educación ambiental que han dado resultados positivos. Es posible obtener buen éxito didáctico trabajando sobre paisajes sencillos y triviales; no hay que buscar necesariamente parques nacionales, ni siquiera lugares sobresalientes. Lo que importa es descubrir la coherencia estructural y funcional de los sistemas y saber interpretar los hechos a medida que se van presentando. El autor recomienda tratar conjuntamente los aspectos científicos y estéticos, no yuxtapuestos caprichosamente, como podrían estar en un plan de estudios académico, sino obedeciendo a la congruencia impuesta por la propia interpretación. Es importante también la comunicación entre distintas culturas, por ejemplo entre la cultura urbana y la cultura rural autóctona; no hay que olvidar que la última puede tener conocimientos empíricos muy profundos sobre su propio entorno. No es bueno ignorarlos o despreciarlos, suplantándolos brutalmente por la aplicación de métodos a veces poco adecuados. (O. B.)

SCIENTIFIC PROGRESS. A Study Concerning the Nature of the Relation Between Successive Scientific Theories, por Craig Dilworth. D. Reidel Publishing Company; Dordrecht (Holanda) 1981; 155 páginas. En los seis primeros apartados del libro, Dilworth critica la concepción de la ciencia y del progreso científico que se da en el empirismo lógico y en Popper, mostrando las insuficiencias del "modelo deductivo" de la ciencia que ambas posturas presuponen, y haciendo explícitas las dificultades con que tropiezan al afrontar algunos problemas tales como el cambio de significado de términos comunes a teorías diferentes, el conflicto entre teorías sucesivas, la naturaleza de las leyes y teorías científicas y la aplicación de las concepciones epistemológicas al desarrollo real de la ciencia. Aunque se trata de cuestiones ampliamente desarrolladas en la epistemología contemporánea, el estudio de Dilworth tiene el mérito de señalar con claridad algunas limitaciones de las posturas mencionadas mediante su reconstrucción en términos del "modelo deductivo" de la explicación científica, y de mostrar, en ese contexto, su incapacidad para dar cuenta de algunos aspectos del progreso científico.

Desde el apartado siete hasta el 12 (y último), Dilworth desarrolla su concepción de la ciencia y del progreso científico, intentando superar los inconvenientes de las posturas que ha criticado. El punto de partida son algunos rasgos de los bien conocidos enfoques de Kuhn y Feyerabend acerca de la "inconmensurabilidad" de las teorías en conflicto (apartado 7). Intenta proseguir esos enfoques de modo natural, para llegar a una concepción en la que la transición de una teoría a otra se basa en consideraciones a la vez racionales y empíricas (p. 52), de modo que se evite la acusación de "irracionalidad"

que frecuentemente se ha dirigido contra los enfoques mencionados. El camino que elige es la exposición del "modelo de la forma" (en el libro, "The Gestalt Model") (apartado 8), del que pasa a continuación hasta su "concepción perspectivista de la ciencia" ("The Perspectivist Conception of Science"), basada en el modelo de la Gestalt, aunque el autor aclara que puede llegarse a ella por otros caminos (apartado 9). El apartado 10 está dedicado al desarrollo de la concepción perspectivista en el contexto de la teoría cinética de los gases; el 12, a su aplicación a las concepciones de Newton, Kepler y Galileo. Entre los dos últimos, el apartado 11 trata acerca de la visión de la ciencia desarrollada por Sneed y Stegmüller ("The Set-Theoretic Conception of Science"), que Dilworth considera insuficiente: no podía faltar este tema, puesto que Stegmüller ha intentado precisamente completar las interpretaciones de Kuhn –coincidiendo en esto con Dilworth–, mediante el formalismo desarrollado por Sneed para la física matemática.

Dilworth aborda el tema del progreso científico estudiando la relación entre teorías científicas sucesivas, o sea, entre una teoría aceptada y otra que pretende superarla. Su tesis central es que se da progreso cuando una teoría es científicamente más aceptable que otra, y que esto ha de juzgarse en función de tres factores: su exactitud o precisión ("accuracy"), el alcance o amplitud de sus aplicaciones ("scope" o "generality") y la sencillez ("simplicity"); los dos primeros factores son empíricos y el tercero tiene naturaleza más racional. La precisión se comprueba mediante los resultados de las medidas en el dominio común en que las dos teorías son aplicables. El mayor alcance de una teoría significa que es aplicable a más problemas concretos. La sencillez depende del número de modificaciones "ad hoc" que hay que introducir en una teoría para explicar los datos empíricos.

La tesis de Dilworth se apoya en una concepción de las teorías científicas según la cual éstas no son –como pretende el "modelo deductivo"– un sistema de enunciados sobre cuya verdad hay que decidir, sino un "modelo aplicado" del que se pueden derivar leyes empíricas. De este modo, Dilworth se aparta decididamente de lo que suele llamarse la "concepción lingüística de las teorías", íntimamente relacionada con una epistemología que se centra en los aspectos lógico-formales, sin llegar a decir –como es frecuente en la epistemología actual– que una teoría sea una "visión del mundo" ("Weltanschauung"). Según Dilworth, las teorías científicas son "perspectivas conceptuales" que se aplican a determinadas situaciones de hecho, y el juicio sobre ellas no versa acerca de su verdad o falsedad, sino acerca de su mayor o menor aplicabilidad. Por otra parte, Dilworth insiste en la distinción entre las teorías y las leyes, considerándola de gran importancia epistemológica, y reprocha a las posturas que critica el hecho de que valoran insuficientemente esa distinción; en cuanto a las leyes científicas, las considera "empíricas" o "experimentales" cuando se llega a ellas mediante operaciones con instrumentos de medida, y "teóricas" cuando se deducen de un modelo teórico. Dilworth examina en detalle muchos otros aspectos de los temas mencionados, y aplica su concepción a los problemas del conflicto entre teorías –que implica una "incompatibilidad perspectiva" y no una contradicción entre enunciados–, del cambio de significado de los térmi-

nos –que queda superado– y de la "inconmensurabilidad" de las teorías sucesivas –que tampoco implica una contradicción entre ellas, sino un cambio de "perspectiva conceptual"–. Es importante señalar que los "modelos" que están en la base de las teorías no tienen por qué ser representables imaginativamente.

Bajo el aspecto filosófico, es positivo que Dilworth insista en el carácter limitado de sus tesis, con las que no pretende solucionar problemas de la filosofía general o de la teoría del conocimiento; de este modo, evita la tentación, en que no pocos autores caen, de extrapolar posturas concretas de la epistemología para justificar indebidamente concepciones filosóficas que las sobrepasan.

Por otra parte, Dilworth afirma justamente que sus tesis no excluyen una postura realista ni una instrumentalista acerca de las teorías científicas (p. 117). En una comunicación personal me justificaba esta importante cuestión del modo siguiente: "Tal como lo veo, un realista exigiría –para que una teoría científica fuera aceptable– la existencia actual de las entidades hipotéticas que postula. Yo no exijo tanto. Todo lo que pediría es que lo descrito por una teoría sea en principio capaz de *concebirse* como una situación *física* posible, y el motivo para formular solamente esta exigencia más débil es, en parte, que puede suceder con frecuencia que no se esté en condiciones de aplicar el criterio realista, ya que no siempre podemos saber si aquello a lo que se refieren nuestros 'términos teóricos' existe o no. Por tanto, no sería correcto decir que estoy en desacuerdo con un realista: simplemente, soy algo más tolerante respecto a los criterios de aceptabilidad de las teorías científicas". Sin duda, esta postura es correcta, en cuanto refleja lo que de hecho sucede en la práctica científica. Sólo hay que añadir que los problemas acerca del sentido realista del conocimiento científico quedan abiertos, y constituyen un serio desafío a la profundidad de la actual filosofía de la ciencia. Ingvar Johansson, a quien Dilworth agradece la lectura y comentarios de su libro, ha escrito que todos los actuales filósofos de la ciencia son solamente figuras de transición, debido a que no plantean de modo satisfactorio la "filosofía de la percepción" (quizá sería mejor decir "filosofía del conocimiento", en sentido realista).

Por lo que respecta a la "concepción perspectivista" de la ciencia, reconociendo su valor, puede dudarse acerca de su generalidad. Dilworth señala con acierto no pocos defectos de la "concepción lingüística", muestra cómo se superan mediante la concepción perspectivista y centra su atención principalmente en los "cambios de perspectiva" que implica la sucesión de teorías, de modo que los enunciados científicos parecen quedar relegados a un segundo término. En realidad, Dilworth advierte que el problema de la ciencia no es la *verdad* de unos enunciados, sino la *aplicación* de conceptos relacionados en un modelo; con esto no se niega la importancia de los enunciados: por ejemplo, como ya se ha señalado, la precisión de los enunciados respecto a las medidas es uno de los tres factores básicos que permiten juzgar si se da el progreso científico. Pero sería deseable un mayor equilibrio entre los dos aspectos de las teorías: la "formulación" de enunciados y el "sistema conceptual" supuesto en una teoría están relacionados, sin confundirse ni oponerse. Al insistir de modo continuo en los "sistemas conceptuales", Dilworth deja en la penumbra las cuestiones relativas a los enunciados y, por tanto, el problema de su verdad o falsedad.

La obra de Dilworth es clara y sugerente. Las tesis básicas quedan expuestas con la elegante sobriedad de la filosofía que puede llamarse científica por su método y no sólo por su temática. Las dos observaciones anteriores, acerca del realismo y de la función de los enunciados –dos temas íntimamente relacionados–, sugieren un ámbito de problemas que podrían ser tratados sin dificultad mediante el enfoque básico de Dilworth, aunque requerirían el complemento de un formalismo más complejo y de una teoría realista del conocimiento desarrollada.

Es interesante señalar, por último, que la postura de Dilworth depende en mucho mayor grado de Campbell que de Kuhn o Feyerabend; puede verse al respecto la cita de Campbell de la página 110. Para el científico, esto tiene una cierta lógica: aunque las obras principales de Campbell datan de los años 20, y tienen las consiguientes limitaciones, se deben a un autor familiarizado con la práctica de la ciencia, y es significativo que a más de 50 años de distancia puedan inspirar un enfoque apto para afrontar no pocos rompecabezas de la moderna filosofía de la ciencia. (M. A.)

INTRODUCCIÓN A LA ANTROPOLOGÍA GENERAL, por Marvin Harris. Madrid; Alianza Universidad Textos; 1981; 599 páginas. En los últimos años puede advertirse una cierta dispersión en la que se fracciona el discurso antropológico a partir de subdivisiones designadas como antropobiología y antropología cultural, y en ésta la etnología, la arqueología y la lingüística; y finalmente, una antropología social que entendida como una rama de la sociología ha iniciado muchas de las confusiones epistemológicas predominantes. Lo cierto es que cada una de dichas subdivisiones acude a otras todavía más internas para explicar sus datos, de manera que puede hablarse de antropología cognitiva, económica, política, ecológica y psicológica, sin que ninguna de éstas se sienta extraña a una concepción antropológica de sus problemas.

El resultado de esta fragmentación lo constituye la existencia de "antropologías" que muchas veces tienen de tales sólo el título, pues si en ellas se acusa el defecto de manifestarse como reducciones circulares de una realidad mayor, la sociedad en su proceso e integración, terminan siendo presentadas como teorías totales. En este punto es donde la circularidad o conjunto referido al análisis de una parcialidad acaba justificando su existencia en el determinismo, sea éste el material (naturaleza, ecología, tecnología, economía), sea el espiritual en sus versiones religiosa y estética, o bien sea que la racionalidad de lo mental o la organización psicológica sean per se temas completos de lo antropológico.

El efecto de esta operación determinista, a que conduce la conversión de lo parcial en teoría total de lo antropológico, no consiste sólo en el hecho de la calidad relativa que se obtiene de la aplicación del análisis intensivo a una realidad parcial, sino que aparece en el hecho de que la reducción de una categoría parcial convertida en categoría total lleva a la indeterminación al resto de la realidad no implicada. En cierta manera, y es lo más significativo, el plano antropológico parcial se aborda, en tal extremo, desde la perspectiva de una impreparación académica en lo que refiere a fuentes de carácter integral, de manera que muchos determinismos cabe achacarlos al principio de la circularidad parcial y al hecho de que la selección de las categorías de análisis responde más a la reducción intelectual del mismo antropólogo que a las mismas cualidades del sistema que se observa y racionaliza. Vale aquí decir que el peor enemigo del determinismo reside en la misma generalidad de su eclecticismo empírico.

Esta mínima consideración viene a cuento de la obra de Marvin Harris, una de las figuras relevantes de la antropología norteamericana contemporánea y cuya producción intelectual nos provee de estímulos críticos, por lo menos en el hecho de reafirmarnos en nuestra convicción de que la antropolo-

gía sigue siendo el arte de atribuir sentido a la realidad humana comparada, y en este extremo este arte se prolonga a la idea de significar como relativismo cultural lo que son representaciones específicas del discurso humano. Por añadidura, y cuando sea también especificidad adaptativa, procedemos a relativizar el mismo comportamiento biológico. Y ambos no los relativizamos únicamente por su especificidad adaptativa, sino que relativizamos el resultante de la explicación determinista. En todo caso, ambas, cultura y biología, tienen en su contexto interno un algo de azar todavía inmedible en términos de su historiedad y de su comprensión. Esta inmedición se relaciona con el carácter cualitativo de las categorías antropológicas y con el hecho del indeterminismo cuando se consideran estas categorías dentro de una totalidad inabarcable a la observación de campo.

La existencia de diferencias intraespecíficas se convierte en impedimento para todo intento de atribuir causas únicas a los fenómenos antropológicos, y aquí es donde el materialismo cultural que proclama Harris se nos aparece como una falacia conceptual. Al leer este libro nos volvemos inevitablemente a las posturas de Alexander Goldenweiser (1937) y a las de Alfred L. Kroeber (1948) quienes, en su respectiva "Antropología", presentaban ésta como una disciplina ocupada en el estudio de la humanidad vista como una historia de sus cambios biológicos y culturales y de sus diferentes adaptaciones actuales en el espacio, esto es, en su diversidad intraespecífica. Harris ha actualizado los datos de este enfoque de Goldenweiser y de Kroeber, y desde luego lo ha hecho ampliando algunos sectores del conocimiento antropológico y reduciendo, en cambio, otros a partir de la relativización personal de las temáticas. Por ejemplo, la tecnología es insumida en el contexto de la energía, de la economía y de la ecología.

Para nosotros, la definición de materialismo cultural hecha por Harris (p. 133) es insuficiente y confusa porque decir, por ejemplo, que "El principio teórico-básico del materialismo cultural es que los modos *etic* y conductuales de producción y reproducción *determinan probabilísticamente* la estructura, la superestructura y todos los componentes *emic* y mentales de los sistemas socioculturales", es tanto como ser determinista sin demostrar la determinación; o sea, sin mostrar la cadena de acontecimientos que siguen a una primera causalidad, en este caso material e infinita. Ni siquiera la declaración del probabilismo estructural constituye una noción eficiente, pues toda producción material no sólo es un modo, sino que es una cualidad variable derivada del conocimiento existente sobre la realidad, de manera que este conocimiento aparece como límite cultural de la realidad material.

En nuestro caso, atribuimos cambios ideológicos a infraestructuras materiales que parten de situaciones diferentes (cristianismo, comunismo, socialismo, liberalismo), tanto como de situaciones semejantes, con lo cual el determinismo pierde su carácter universal. Aquí es más válido significar los determinismos en el ámbito de los estadios primarios de cultura, cuando la acción de la naturaleza sobre los comportamientos está menos controlada por el hombre. Pero en la medida en que el dominio de la cultura empieza a condicionar el desarrollo mismo de la naturaleza, aparece el dominio de la superestructura. En este punto, el determinismo del materialismo cultural convierte el efecto en causa y provoca una confusión epistemológica insoluble: la que resulta de unir la materialidad cultural, puro instrumento de la acción, con la acción misma partiendo de adaptaciones emanadas desde la superestructura. En todo caso, la construcción de teorías antropológicas dentro del marco determinista representa una construcción inmedible y acientífica.

En todo caso, aquí uno puede indicar que cuanto más compleja es una estructura, mayor es su grado de variabilidad causal interna, de manera que si los recursos de la infraestructura son distribuidos desigualmente, también entre materiales iguales (los seres humanos) son diferentes los recursos intelectuales y cognitivos, así como las proyecciones simbólicas y psicológicas derivadas de su aplicación.

Harris no acaba de situar su alternativa entre lo social y lo cultural. No nos parece acertado cuando dice que la arqueología es una ciencia social (p. 136), cuando en realidad es obvio que su énfasis es estrictamente cultural al depender del estudio de producciones materiales. Por lo menos, el mismo Harris no parece advertir que los materiales arqueológicos corresponden a categorías culturales (¿estrictamente éticas?) en sí asociales.

En cuanto a que los amerindios alcanzaron en América el estatus de *Homo sapiens* (p. 173), cabe decir más bien que ya lo traían consigo cuando se aposentaron en dicho continente. Parecería como si la causa de haber alcanzado a ser *sapiens* la debieran a su adaptación posterior. No es consecuente con el principio del materialismo cultural cuando atribuye el "ritmo lento" o retraso cultural de los Andes prehispánicos respecto de otros puntos geográficos "a las diferentes dotaciones naturales de las regiones nucleares americanas y del Oriente Medio" (p. 189). Si fuera consecuente, el materialismo cultural daría otra explicación: la de los medios empleados, incluidos los históricos. Igualmente, trabaja sobre conjeturas cuando dice: "si se les hubiera dejado solos a los indios americanos, éstos también habrían acabado por descubrir las superiores cualidades del hierro y el acero" (p. 190).

Podemos añadir al contexto falaz de este materialismo cultural la idea del papel decisivo de la infraestructura sin más, una infraestructura que consiste en ser lo material una fuerza gobernada desde sí misma, sin pensar que la cultura tiene como condición previa de su expresión la existencia inteligente.

Lo cierto es que cuanto más compleja y avanzada es una cultura, mayor es el papel del programa, del laboratorio y de la prefiguración de la realidad material hasta convertirse ésta en sujeto de manipulación por parte de la superestructura. El papel de las élites dirigentes en las sociedades avanzadas, o sea, políticos, intelectuales, líderes, científicos, empresarios, religiosos y artistas es de tal envergadura que es a partir de sus decisiones como se diseñan las condiciones en que actuará la producción material. La transformación del medio material es un producto inteligente que actúa por prefiguración de la realidad.

En realidad, esta "Introducción a la Antropología General" cubre un amplio campo de problemas, y en nuestra opinión su mayor excelencia reside en el carácter integrador que adopta cuando se ocupa de las diferentes disciplinas que constituyen la antropología. En tal caso, se trata de un libro excelente, bien escrito y planificado. Es en el orden teórico donde nos hemos sentido inclinados a impugnar lo que entendemos es sólo un enunciado, el del materialismo cultural, sin verificación controlada. Pero hay otras observaciones a considerar. Una de ellas es la de que el capítulo 23, referido a *Lenguaje y Cultura*, debió haber sido colocado tras el biológico, o por lo menos inmediato al que aparece dedicado a *La Cultura*. Y por otra parte, esta antropología de Harris produce en nosotros la sensación de ser autosuficiente en un sentido: en el de que toda la bibliografía es de habla inglesa, con lo cual parece ignorarse que en los países de habla española, francesa y alemana existen magníficas antropologías. (C. E.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

ESPERANZA DE VIDA Y CRECIMIENTO DEMOGRAFICO EN EL TERCER MUNDO

THE POLICY IMPLICATIONS OF FIELD EXPERIMENTS IN PRIMARY HEALTH AND NUTRITION CARE. Davidson R. Gwatkin, Janet R. Wilcox y Joe D. Wray en *Social Science & Medicine*, vol. 14C, n.º 2, págs. 121-128; 1980.

HOW MANY DIE? A SET OF DEMOGRAPHIC ESTIMATES OF THE ANNUAL NUMBER OF INFANT AND CHILD DEATHS IN THE WORLD. Davidson R. Gwatkin en *American Journal of Public Health*, vol. 70, n.º 12, págs. 1286-1289; diciembre, 1980.

INDICATIONS OF CHANGE IN DEVELOPING COUNTRY MORTALITY TRENDS: THE END OF AN ERA? Davidson R. Gwatkin en *Population and Development Review*, vol. 6, n.º 4, págs. 615-644; diciembre, 1980.

QUARKONIO

TEORÍAS GAUGE DE LAS FUERZAS ENTRE PARTÍCULAS ELEMENTALES. Gerard 't Hooft en *Investigación y Ciencia*, n.º 47, págs. 58-77; agosto, 1980.

RECENT EXPERIMENTAL ADVANCES IN POSITRONIUM RESEARCH. Arthur Rich en *Reviews of Modern Physics*, vol. 53, n.º 1, p. 127-165; enero, 1981.

OBSERVATION OF AN η'_c CANDIDATE STATE WITH MASS 3592 ± 5 MEV. C. Edwards et al. en *Physical Review Letters*, vol. 48, n.º 2, págs. 70-73; 11 de enero de 1982.

IMAGENES POR RMN EN MEDICINA

NUCLEAR MAGNETIC RESONANCE IMAGING IN MEDICINE. Dirigido por L. Kaufman, L. E. Crooks y A. R. Margulis. Igaku-Shoin, Ltd., Tokyo, 1981.

PROCEEDINGS OF AN INTERNATIONAL SYMPOSIUM ON NMR IMAGING, BOWMAN GRAY SCHOOL OF MEDICINE WINSTON-SALEM, N. C., OCTOBER 1-3, 1981. Dirigido por N. Karstaedt, R. L. Witcofski y C. L. Partain. Bowman Gray School of Medicine Press, 1982.

PRINCIPLES OF NMR IMAGING. I. L. Pykett, J. H. Newhouse, F. S. Buonanno, T. J. Brady, M. R. Goldman, J. P. Kistler y G. M. Pohost en *Radiology*, vol. 143, págs. 157-168; abril, 1982.

BASES GENETICAS DE LA DIVERSIDAD DE ANTICUERPOS

EVIDENCE FOR SOMATIC REARRANGEMENT OF IMMUNOGLOBULIN GENES CODING FOR VARIABLE AND CONSTANT REGIONS. Nobumichi Hozumi y Susumu Tonegawa en *Proceedings of the National Academy of Sciences of the United States of America*, vol. 73, n.º 10, págs. 3628-3632; octubre, 1976.

SEQUENCES OF FIVE POTENTIAL RECOMBINATION SITES ENCODED CLOSE TO AN IMMUNOGLOBULIN κ CONSTANT REGION GENE. Edward E. Max, J. G. Seidman y Philip Leder en *Proceedings of the National Academy of Sciences of the United States of America,* vol. 76, n.º 7, págs. 3450-3454; julio, 1979.

THE ORGANIZATION OF IMMUNOBLOGULIN GENES AND THE ORIGIN OF THEIR DIVERSITY. P. Leder, E. E. Max y J. G. Seidman en *Immunology 80*, dirigido por M. Fougereau y J. Dausset. Academic Press, Inc., 1980.

LA ASIMETRIA DE LAS PLATIJAS

ON THE BIOLOGICAL BASIS OF HUMAN LATERALITY, I: EVIDENCE FOR A MATURATIONAL LEFT-RIGHT GRADIENT. II: THE MECHANISMS OF INHERITANCE. M. C. Corballis y M. J. Morgan en *The Behavioral and Brain Sciences*, vol. 2, págs. 261-336; 1978.

FLATFISHES AND THE INHERITANCE OF ASYMMETRIES. David Policansky en *The Behavioral and Brain Sciences*, en prensa.

CHORROS COSMICOS

EL CURIOSO ESPECTRO DE SS 433. Bruce Margon en *Investigación y Ciencia*, número 51, págs. 22-36; diciembre, 1980.

RELATIVISTIC JET PRODUCTION AND PROPAGATION IN ACTIVE GALAXIES. M. J. Rees, M. C. Begelman y R. D. Blandford en *Annals of the New York Academy of Sciences*, en prensa.

EL CALENDARIO GREGORIANO

PAPI ASTRONOMI EPATTE: LUIGI GIGLIO, ASTRONOMO "PRIMES AUCTOR" DELLA RIFORMA GREGORIANA DEL CALENDRIO. Giulio Aromolo. Istituto della Stampa, Nápoles, 1963.

FURTHER ADJUSTMENT OF THE GREGORIAN CALENDAR YEAR. E. R. Hope en *The Journal of the Royal Astronomical Society of Canada*, 1.ª parte, vol. 58, n.º 1, págs. 3-9, febrero, 1964; 2.ª parte, vol. 58, n.º 2, págs. 79-87, abril, 1964.

DIE GREGORIANISCHE KALENDERREFORM IM URTEIL ZEITGENÖSSISCHER ASTRONOMEN. Peter Aufgebauer, en *Die Sterne*, vol. 45, n.ºs 4/5, págs. 118-121; 1969.

EL CUERPO CAROTIDEO

A DISCUSSION OF POSSIBLE TRANSMITTER OR GENERATOR SUBSTANCES IN CAROTID BODY CHEMORECEPTORS. C. Eyzaguirre y P. Zapata en *Arterial Chemoreceptors*, (págs. 231-251), dirigido por R. W. Torrance. Blackwell. Oxford, 1968.

THE PERIPHERAL ARTERIAL CHEMORECEPTORS. Dirigido por M. J. Purves. Cambridge University Press, London; 1975.

CHEMORECEPTION IN THE CAROTID BODY. Dirigido por H. Acker S. Fidone, D. Pallot, C. Eyzaguirre, D. W. Lübbers y R. W. Torrance. Springer Verlag, Berlín; 1977.

MODULATORY ROLE OF DOPAMINE ON ARTERIAL CHEMORECEPTORS. P. Zapata en *Non-S.riatal Dopaminergic Neurons*, dirigido por E. Costa y G. L. Gessa. Raven Press, Nueva York; 1977, págs. 291-298.

A STUDY OF SOME MECHANISMS OPERATING IN CAROTID BODY CHEMORECEPTORS. C. Eyzaguirre y S. J. Fidone en *Studies in Neurophysiology*, págs. 109-129, dirigido por R. Porter. Cambridge University Press, London; 1978.

TEMAS METAMAGICOS

COSMIC VIEW: THE UNIVERSE IN FORTY JUMPS. Kees Boeke. John Day Company, 1957.

THE LORE OF LARGE NUMBERS. Philip J. Davis. Random House, Inc., Nueva York; 1961.

TALLER Y LABORATORIO

THE PERCEPTION OF TWO PURE TONES. John S. Rigden en *Physics and the Sound of Music*. John Wiley & Sons. Inc., 1977.